ce Library

JUL 12 1946

# JOURNAL OF CONSULTING PSYCHOLOGY

VOL. X, NO. 2
MAR.–APR.
1 9 4 6

## CONTENTS

AMERICAN SOCIAL PSYCHOLOGY AND THE WAR — DORWIN CARTWRIGHT — 67

MILITARY TRAINING AND LEARNING THEORY — DAEL WOLFLE — 73

ASSESSMENT OF OSS PERSONNEL — HENRY A. MURRAY AND DONALD W. MacKINNON — 76

DEVELOPMENTS IN OCCUPATIONAL CLASSIFICATION — CARROLL L. SHARTLE — 81

INDUSTRIAL IMPLICATIONS OF WARTIME DEVELOPMENTS IN PSYCHOLOGY — MORRIS S. VITELES — 85

VALIDITY FOR WHAT — JOHN G. JENKINS — 93

MINNESOTA MULTIPLASTIC PERSONALITY INVENTORY FOR GROUP ADMINISTRATION — ARTHUR BURTON AND CHARLES J. BRIGHT — 99

PROFESSIONAL STATUS AND TRAINING OF PSYCHOLOGISTS. REPORT — LOUIS LONG — 104

ON CERTAIN MISUSES OF THE WECHSLER-BELLEVUE SCALES — A. S. LUCHINS — 109

BOOK REVIEWS: HUMAN RELATIONS IN INDUSTRY — WHAT PEOPLE ARE — YOUNG MAN, YOU ARE NORMAL — PERMANENT LEARNING — 112

AMERICAN : PSYCHOLOGICAL : ASSOCIATION : INCORPORATED

VOLUME X MARCH-APRIL, 1946 NUMBER 2

# JOURNAL OF CONSULTING PSYCHOLOGY

PUBLISHED BY THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

EDITOR

MRS. J. P. SYMONDS, 525 West 120 Street, New York 27, N. Y.

ASSOCIATE EDITORS

HAROLD E. BURTT, PH.D., The Ohio State University
HAROLD E. JONES, PH.D., University of California, Berkeley
ARTHUR W. KORNHAUSER, PH.D., University of Chicago
BERTHA M. LUCKEY, PH.D., Board of Education, Cleveland, O.
CARL R. ROGERS, PH.D., University of Chicago
GOODWIN WATSON, PH.D., Teachers College, Columbia University
F. L. WELLS, PH.D., Department of Hygiene, Harvard University

The JOURNAL OF CONSULTING PSYCHOLOGY was first published as the official journal of the American Association for Applied Psychology, Inc., which was organized in 1937 as an amalgamation of the Association of Consulting Psychologists, Inc., and the Clinical Section, American Psychological Association, Inc., with the affiliation of regional associations for applied psychology. In September, 1945 the American Association for Applied Psychology and the American Psychological Association were reconstituted as one organization under the name of the American Psychological Association, and publication of the JOURNAL OF CONSULTING PSYCHOLOGY continued as one of its official publications.

The JOURNAL OF CONSULTING PSYCHOLOGY publishes results of research bearing on the application of psychology, short articles and communications relating to professional psychological service, reports of the activities of organized groups of professional psychologists with representation in the American Psychological Association, information and notices of interest to psychologists, resources available to the profession, and critical reviews of the professional literature. Contributions are invited from all fields which are significantly related to applied psychology.

### EDITORIAL OFFICE

Manuscripts, general correspondence and publications intended for notice or review should be sent to Mrs. J. P. Symonds, *Editor*, 525 West 120th Street, New York 27, N. Y. Manuscripts are accepted only upon recommendation by the editorial staff. Proof will be submitted to the author at the time of publication. Corrections made by the author are chargeable to him. Manuscripts should be submitted in duplicate. References to publications should conform to the style of this JOURNAL, giving in order: name of author, title of article, full name of periodical, year, volume, number, pages. References to books should give in order: author's name, title of book, place of publication, name of publisher, year, pages. Articles in this JOURNAL are indexed in *Psychological Abstracts*.

### BUSINESS OFFICE

DAEL WOLFLE, PH.D., *Executive Secretary* of the American Psychological Association, Incorporated, and *Business Manager* of the JOURNAL OF CONSULTING PSYCHOLOGY, 1227 Nineteenth Street, N. W., Washington 6, D. C.

Correspondence regarding subscriptions and business matters should be addressed to the publishing office at Colorado Springs, Colorado, or to DAEL WOLFLE, PH.D., *Business Manager*, 1227 Nineteenth Street, N. W., Washington 6, D. C.

The JOURNAL OF CONSULTING PSYCHOLOGY is published bimonthly six times a year. All members of the American Psychological Association who were formerly members of the American Association for Applied Psychology are entitled to receive the JOURNAL upon payment of their annual dues, of which $2.50 is applied to the JOURNAL. Affiliated state or regional associations may arrange a group subscription with the Business Office through their Representative in the Board of Affiliates whereby members of the group subscription are entitled to receive the JOURNAL upon payment of $2.50 added to the dues of their association. The individual subscription is $3.00 a year and single copies are 60 cents a number. Postage is prepaid. Claims for nonreceipt of a number will be honored in the Business Office within two weeks after receipt of the next succeeding number.

Copyright 1946 by American Psychological Association, Inc. Printed at 23 West Colorado Avenue, Colorado Springs, Colo., and entered as second-class matter at the Colorado Springs, Colo., post office.

Volume X MARCH-APRIL, 1946 Number 2

JOURNAL OF
CONSULTING PSYCHOLOGY

# AMERICAN SOCIAL PSYCHOLOGY AND THE WAR[1]

BY DORWIN CARTWRIGHT
RESEARCH CENTER FOR GROUP DYNAMICS
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

WORLD WAR II was total war, involving the participation of civilian populations as never before. It required in this country the rapid mobilization of a tremendous military machine from a civilian population reluctant to go to war. To develop overwhelming military power, a task of social organization was required unparalleled in history. Practical problems of social engineering sprang up over night which required solution before lunch. Many of these problems were tackled by social psychologists; many by psychologists who had not thought of themselves as expert in this field; many were undertaken by social scientists from other academic disciplines; many by men of practical affairs; and, unfortunately, many were never faced at all. In any event, social psychologists did make major contributions to the war effort, and in return social psychology comes out of the war greatly altered by the experience.

[1] A description of progress in social psychology in other countries cannot be attempted here. Considerable work has been done, however, particularly in England, and deserves special treatment.

Prior to the war it could not be said exactly that social psychologists were cloistered in an ivory tower, for their subject-matter had to do with social events and behavior. Yet, the empirical study of significant social behavior had barely begun, and the center of gravity of research was still the investigation of attitudes, particularly those of college students or others readily accessible to the campus. To be sure, public opinion polls were in existence, and field studies were occasionally done, but only during the war did a phrase like "action research" come into being and only very recently have social psychologists spoken with much confidence of the possibility of a scientifically grounded social engineering. Before the war a social psychologist could stay near the campus with a clear conscience; now such conduct is, to say the least, out of date.

A major prewar deterrent to the study of "real" social behavior lay in the fact that rarely did social psychologists have natural access to social groups nor sufficient money to carry out research on the grand scale required in the study of major social problems. The war provided both of these prerequi-

sites. Whatever happens to specific wartime governmental social research agencies and whatever the fate of the social sciences in the National Science Bill, governmental expenditures for social psychological research will never again reach the low levels of the prewar era. Government administrators in sufficiently large numbers have come to see the value of such research to guarantee its wide-scale continuation. Moreover, by working with government, business, labor, and social agencies during the war, social psychologists have developed effective contacts with practical people so that the larger social world is now more readily accessible as a scientific laboratory.

In yet another way the war has changed the whole manner in which social psychologists approach their work. Concentration on administrative rather than theoretical problems has thrown social psychologists into intimate contact with neighboring academic disciplines, particularly sociology, anthropology, and economics. In the field of public opinion research, for example, all four fields were represented in equally central positions in the research groups in Washington. It has become clear that practical social problems require the combined knowledge and skill of various disciplines if they are to be fruitfully solved, and many now believe that even theoretical and systematic problems require the same type of interaction.

It is one thing, however, to pool knowledge, but it is an entirely different matter to work *cooperatively* upon the *same* project. Before the war, the whole emphasis of our graduate schools was the training of individuals to be competent, clever, and critical. To be able to find flaws in others' research gave prestige. This emphasis upon competitive individualism was meager training for the cooperative work required by large wartime research organizations, and many highly competent people found the social adjustments difficult to make. But those who learned how to take full advantage of the potentialities of real cooperative research now believe that the goal is worth training for. It is possible that graduate students of the future may be given some training in real cooperation, and it is even barely conceivable that in the dim future a group of students may be permitted to submit a joint thesis.

A final feature of the war experience promises less profound but still important effects upon future work in social psychology. Most wartime research was done under terrific pressure of time. At first academic psychologists felt deeply frustrated that the full significance of data could not be extracted before dead lines imposed by impatient administrators. In time, however, many felt that working under pressure forced upon them a new skill in distinguishing essentials from inconsequential by-paths. While it is possible that the rate of production of these people may decline as the external pressures are reduced, it seems likely that research efficiency and productivity will be valued more highly than ever before.

These are the gross features which characterize social psychology as it emerges from the war. But let us be more specific. What was the exact nature of the activities that brought about this change? What did social psychologists do to change their outlook so basically?

## MAJOR AREAS OF ACTIVITY

The great variety of research activities undertaken by social psychologists defies ready classification. It would be possible to group research by organizations or by techniques. In the brief

space available here, however, many important projects would have to be slighted. On the other hand, it does appear feasible to enumerate the major purposes served by the various types of projects and to sight only illustrative examples under each.

1. Civilian Morale and Welfare

The problems of maintaining civilian morale, perhaps more than any others, captured the interest of social psychologists during the war. The Society for the Psychological Study of Social Issues published its yearbook[2] on the subject early in 1942; the Harvard Seminar in Psychological Problems of Morale prepared worksheets which were widely circulated among governmental agencies; the Bureau of Intelligence of O.W.I., through both intensive and polling interviews, made a number of national studies; and private polling organizations investigated many aspects of the problem for various agencies. Other research groups developed morale scales and investigated the more theoretical problems.

Particular attention was given, principally by nongovernmental groups, to reducing the disruptive effects on morale caused by rumor. The development of "rumor clinics" in which newspapers publish rumors branded as such came early in the war following the work of Floyd and Gordon Allport and of Knapp. Interest in rumor waned, however, as victory became more certain, and the few projects undertaken by the government were soon abandoned.

Even though Hitler's program of "divide and conquer" made it all too clear how disastrous to national unity friction between groups can be, only a relatively small program of research on minority problems was undertaken by governmental agencies. One notable project, however, was that of the Army concerning the effects of various types and degrees of segregation of troops. Private research groups did the bulk of the work, a particularly significant development being the establishment of an "action research" organization, the Commission on Community Interrelations of the American Jewish Congress. Social psychologists served more directly in the line of action through various community agencies and through educational projects such as the Harvard group's *ABC's of Scapegoating*.[3]

Within the general area of civilian morale a particularly crucial wartime problem was the maintenance of morale among industrial workers. Since it was clear that military power was directly dependent upon such phenomena as absenteeism, turnover, conflict between old and new workers, and the adjustment of minority groups of workers, various governmental and private groups embarked upon extensive research and action programs. Intensive interviews with representative samples of workers revealed the level of morale in the major industries; case studies of individual plants added further knowledge; while evaluation of specific morale-building devices, such as awards, movies, music, and the like helped guide the total program. Basic research upon the effects of organization and atmosphere of groups by McGregor, Lippitt, Bavelas, French, and others led to extensive leadership-training programs in industry, government, and the Armed Forces.

The importance of the media of communication for maintaining morale in time of war was not neglected by social psychologists. Prior to the war basic research by Lazarsfeld, Lasswell, and

[2] G. Watson, editor. *Civilian Morale*. Boston: Houghton Mifflin, 1942.

[3] Harvard University, Department of Psychology. *The ABC's of Scapegoating*. Chicago: Central YMCA College, 1942.

others, had prepared the ground for special wartime projects. In the motion picture industry social scientists have worked closely with the Hollywood Writers Mobilization and others to maximize the contribution of movies to winning the war. The establishment of the *Hollywood Quarterly* is a postwar offspring of this marriage. A specific application of research on newspaper reading was F. H. Allport's study of the effect of war news headlines upon morale which led to a change of policy in the release of news by the Navy.

Bearing more directly on civilian welfare was the study of the wartime problems of children. Although little work was undertaken by governmental agencies, social psychologists were active. Through studies of family disruption caused by the draft, of neglected children of working mothers, of war-created anxieties, and of other tragic impacts of war on children, much was learned about the social and motivational make-up of the child.

2. Psychological Warfare and Military Administration.

This war witnessed extensive expansion of what is known as psychological warfare, including such techniques as short-wave propaganda, dropping leaflets or planting rumors behind enemy lines, and informing the enemy that specific cities would be bombed. While much of the work of the O.S.S., the Psychological Warfare Division of the Army, the O.W.I., and other agencies, is still secret, it is possible to indicate the nature of some of the activities of this kind.

One of the first war agencies to employ social psychologists was the Foreign Broadcast Intelligence Service under the research supervision of Goodwin Watson. Through systematic monitoring and analysis of broadcasts from foreign countries, regular reports on the morale and intentions of the enemy could be made available to key governmental officials. This work aided materially in the preparation of propaganda directed against the enemy, another activity in which social psychologists played central roles.

In tackling the problem of how to make individuals or units surrender more readily, various types of research were employed. For example, captured letters and diaries were analyzed to reveal weak spots in the enemy's psychological armor. Studies of the cultural peculiarities of different groups, sometimes employing actual experiments with members of different nationality groups, proved most helpful.

The whole question of national character was forced to the top of research programs by these and other demands of psychological warfare and military operation in foreign lands. The preparation of propaganda campaigns against the enemy, maintenance of friendly relations with allies, administration of occupied territories, and many other tasks of waging war could not be effective if carried out in a vacuum of cultural understanding. Throughout the war, anthropologists and social psychologists worked closely together on these problems.

A monumental job of evaluating national morale in enemy countries was carried through by the Morale Division of the Strategic Bombing Survey under the direction of Rensis Likert. Working both in Germany and Japan, a large staff of social scientists systematically appraised the effects of strategic bombing on the will of the enemy to resist. Both from the point of view of techniques and findings was this project a milestone in the history of social psychology.

Psychological warfare must defend as well as attack, and the maintenance of high morale among our own troops was guided by more scientifically oriented procedures in this war than ever before. Outstanding were the achievements of the Information and Education Branch of the War Department. Through interviews, questionnaires, and experiments, an elaborate program of morale building was constantly evaluated and improved. In the Army, Navy, and Public Health Service, techniques of leadership training developed in group work and industry were applied to improve the leadership component of good morale.

In yet another way social psychologists contributed to military success. Early in the war it was found that military operations required the rapid organization of liberated and conquered peoples. Thus, in the Italian campaign and later in France, the Army made use of public opinion techniques to obtain important intelligence regarding problems of civil administration. Growing out of this experience, occupation authorities have conducted public opinion polls periodically in Germany since the end of the war. In attempting to shape an intelligent re-education program special studies on the problems of re-educating German youth and prisoners of war have been made. On the whole, however, governing officials have been slower than officials in this country to recognize the assistance which existing techniques can give in their work.

### 3. Civilian Governmental Administration

The use of public opinion research in governmental administration has experienced phenomenal growth during the war. The Department of Agriculture pioneered in this field before the war through its Division of Program Surveys, and with the advent of war, expanded this group to work in urban and rural areas for other agencies. In addition, a polling organization was formed in the O.W.I., and several private polling agencies were employed to do government work.

Agencies dealing with the production and distribution of food faced many problems during the war which called for special research. Anthropologists and social psychologists working together on the Committee on Food Habits of the National Research Council and elsewhere, provided basic knowledge about the dynamic of food habits. More specifically, to evaluate the effectiveness of nutrition programs and measures for increasing agricultural production, the Division of Program Surveys conducted a series of surveys for the War Food Administration.

Bringing together economists and social psychologists, a new field of research has developed within the government to cope with the economic dislocations created by the war. The War Bond Program, from its beginning, was guided constantly by research of the type developed by Likert and his associates. The Office of Civilian Requirements of the W.P.B. learned through periodic national studies exactly what shortages were causing greatest hardships, thereby making it possible to assign priorities for civilian production in an intelligent way. Investigation of the determinants of spending and saving has helped the Treasury, Federal Reserve Board, and O.P.A. to anticipate inflationary developments and to devise combative programs.

In the field of international relations, too, psychologists have been active. The State Department has made regular use of public opinion data, many of which have been collected specifically for its use. While the details of these projects

are still not public knowledge, it is known that information about American attitudes toward international problems has been used in discussions between American officials and those of other members of the Big Three.

## PROSPECT

Many of the research projects undertaken during the war have significance for the postwar period, while others are now possible only because of the war experience. Research upon the problems of readjustment faced by veterans, for example, can be more effective due to spadework done during the war. Postwar planning for the economic future of the country can now be based upon a rapidly developing field of knowledge. The postwar importance of social science in practical affairs is perhaps indicated by the fact that one bill before Congress now proposes to require surveys of consumer demand, labor force, and other psychological-economic matters before allotting annual Federal expenditures. It is to be hoped that social psychological research and knowlege will also be brought to bear upon the significant problems of world organization and international peace.

What the effects of the war experience will be upon departmental organization within the universities is a question for the future, but already signs are pointing to significant trends. Closer collaboration between traditionally separated disciplines is one. At Harvard, for instance, a new Department of Social Relations has been established to be composed of sociologists, social psychologists, and cultural anthropologists. And the older idea of an Institute for Human Relations seems to be growing in popularity. Other indications point to a spread of research organizations working along inter-disciplinary lines upon a common practical problem. Examples are the Industrial Relations Section at M. I. T., which is composed of economists and social psychologists, and the University of Chicago project on the social implications of the atomic bomb.

Provision within the universities of facilities for active work by scientists in the arena of practical affairs is a second trend. At M. I. T., for example, the newly established Research Center for Group Dynamics is organized to develop basic concepts and techniques for studying group dynamics in industry, education, government, and community life. A central fund of methods and theoretical findings will be developed from and applied to these fields.

American social psychology has truly come of age during this war. It has proved itself capable of giving real aid to the country in time of crisis while strengthening itself for future contributions. From our experiences of World War II we now have some reason to hope that the vastly more complicated social problems of the atomic age may be met with sufficiently advanced scientific techniques to allow mankind to survive the social crisis which we face.

# MILITARY TRAINING AND THE USEFUL PARTS OF LEARNING THEORY

BY DAEL WOLFLE
APPLIED PSYCHOLOGY PANEL
NATIONAL DEFENSE RESEARCH COMMITTEE

A STRIKING fact about the successful application of psychological principles to the training of military specialists in World War II was that very substantial help was given by systematically applying only a very few of the principles of learning. Some learning principles are irrelevant to the problems of military training. Others were already well accepted parts of military training doctrine. Between these extremes lie a half dozen or so important principles which have frequently been neglected. By altering training programs and methods to comply with these few principles, the military psychologist was able to improve materially the training of radio operators, gunners, bombardiers, and other military specialists.

Many of the topics traditionally included in a textbook on the psychology of learning are of little practical value in military situations. Opportunities are rare to improve military training by applying our knowledge of the relation between learning ability and chronological age, or by what is known about nonsense syllables, remote associations, or serial relations. The reasons for the uselessness of some of these topics are obvious. In an age range restricted to men at the peak of physical vitality there is little room for any relation between age and learning ability to appear or to be important. Such topics as nonsense syllables, remote associations, and serial relations are unimportant in military training, for memorizing word lists is seldom required.

A number of important principles of learning are already well entrenched in military training doctrine. The usefulness of overlearning was recognized by many officers, although its application varied. The speed required in training men frequently prevented even a close approach to overlearning. In other cases, the enthusiasm of an officer or the necessity of keeping men busy produced an overdose of the principle. A second generally recognized point is that skill is lost during periods without practice. Review and refresher courses were established to prevent or repair this loss. A third principle is the policy of making training situations as much like combat situations as possible. This principle has appealed to the military authorities with the result that men went through some very realistic battle practice and combat maneuvers. Finally, military thinking long ago accepted the doctrine that one learns by doing, and systematic drill has long been required of all recruits.

The principles which were most immediately useful in helping to improve military training, but which had not in 1942 become standard parts of military doctrine, were those which improved the organization and presentation of the material to be learned, and hence improved the motivation of the men in training.

1. *Distribution of practice.*—In war

time men must be trained quickly. There is an understandable temptation to crowd practice into a few days or weeks, and to give the men long hours of drill in order to develop quickly the skills necessary for combat. But this massing of practice is not always necessary. It was sometimes possible, by rearranging a training schedule, to spread practice on an important skill over a greater number of days without increasing the total length of the training period. Neither are the long hours of drill always necessary. It was sometimes possible to reduce the number of hours of drill without loss in the amount learned. In one army camp, men being trained as radio operators drilled seven hours a day on radio code. Reducing the schedule from seven to four hours a day allowed the men to learn just as rapidly, freed time for other activities, and greatly improved the morale of the men in training.

2. *Active participation.*—Active participation by the learner has generally been found superior to passive receptivity. In instruction with which they had a good deal of prewar experience the Services generally made use of this principle. It was sometimes forgotten, however, in the hurried preparation of wartime courses. The military psychologist frequently was able to improve these courses by making them more practical, less bookish, and by insisting upon a greater amount of active participation by the men. One simple application of the principle was in teaching men the phonetic alphabet (Able for A, Baker for B, Charlie for C, etc.). Having the men call out the phonetic equivalents, each time the letters were presented, was shown to be much superior to letting the men see the letter and its phonetic equivalent for an equal number of times.

3. *Variation of material.*—Failure to vary the training material sufficiently has been an easily corrected fault. Some radio schools, for example, repeated the same drill material so frequently that the men partially or completely memorized a particular sequence of characters instead of mastering the individual characters well enough to recognize them in any sequence. The remedy was obvious, and radio code training was improved in a number of schools by the relatively simple method of increasing the amount and variety of practice material given the learner.

The same general principle was applied, with beneficial results, to other types of military drill. A group of radar operators, for example, was found being given such monotonous and unvaried instruction that they became less accurate in locating targets than they were at the start of training. Varying the drills, and giving the men better knowledge of the results of their practice made them improve with practice instead of getting worse.

Greater variation in drill material also helped to improve motivation. Requiring men to perform the same task in the same way day after day can quickly lead to complete boredom. Psychologists, knowing that active and interested participation by the learner speeds learning, were able to suggest new problems, variations on old methods, and other ways of varying the activity of the men in training in order to maintain their interest and enthusiasm.

4. *Accurate records of progress.*—Accurate progress records are useful in two ways. They aid the instructor in coaching the trainee to mastery of his task. And they provide better motivation for the men in training. Psychologists helped to secure accurate records of performance in a variety of ways—by developing simple and usable forms for recording a man's performance, by

improving the kinds of scores which were recorded, and by designing and building scoring attachments for training equipment which was originally built without any. In all these cases the improved scores provided better measures of progress and better checks on the quality of training.

Improved performance records also helped to motivate the men. Posting scores, keeping individual learning curves, and competition between groups or individuals are all good motivational devices. The effectiveness of each depends upon accurate performance records. By improving scoring systems, and by making better use of scores, the psychologist contributed to more effective military training.

5. *Knowledge of results.*—Giving the men in training an immediate knowledge of the results of their trials was one of the most useful single contributions the psychologist made to military training. Instances occurred where men, presumably learning to operate a complex weapon, were simply told to "practice for a while," without being given any information on the results of their practice or the quality of their performance. In some of these cases the men got worse and worse instead of improving with practice. Again the remedy is obvious to one who knows the literature of learning studies. Inform each man, as soon as possible after each trial, how he has done. If well, tell him so. If wrong, tell him where.

As an example, in training gun pointers, tracking errors were reduced by equipping the gun director with an auxiliary sight. Through this "check sight" an instructor saw exactly what the tracker saw, and could tell him immediately when he was tracking well and what mistakes he was making when his tracking was poor.

6. *Systematic lesson plans.*—Putting these principles into practice required the training of instructors and the writing of lesson plans. Good instruction can be obtained from relatively inexperienced teachers if they are told in detail how to proceed. Industrial psychologists and some verbally facile "apparatus men" wrote effective, detailed, and vivid plans in which the material to be learned was organized in simple logical form. Lesson plans of this type state the objective of each lesson and the methods of reaching it, present the simpler material first, the complex later. They tell the instructor when and how to use questions, reviews, examinations, and appropriate teaching aids. By doing these things they solve many of the problems which puzzle inexperienced teachers. They make for better standardized instruction. And they give every trainee some benefit of the knowledge and skill of the expert who wrote the lesson plans.

Numerically this list is not very imposing. Practically, the important stock in trade of the psychologist working on military training problems consisted of these principles:

1. Improve the distribution of practice.
2. Secure active participation of the trainee.
3. Vary the practice material.
4. Develop accurate performance records.
5. Give the men an immediate knowledge of the results of their practice.
6. Write clear detailed plans for the instructor.

The judicious application of these principles helped speed the training of a victorious army and navy.

# ASSESSMENT OF OSS PERSONNEL

By HENRY A. MURRAY
MEDICAL CORPS, A.U.S.
AND
DONALD W. MacKINNON
OFFICE OF STRATEGIC SERVICES

THE TASK of the assessment board of the Office of Strategic Services was that of judging the suitability of each candidate for a proposed assignment overseas, and of placing him in one of five categories: *not recommended; doubtful; recommended with qualifications; recommended; highly recommended.* During the early months of the program the candidate's personality was assessed solely in relation to the mission selected for him by the recruiting and branch officers. Later, after it was discovered that many men, on arrival in the overseas theater of operations, were assigned jobs for which they had not been assessed, an additional practice was instituted: that of assessing a candidate's fitness in relation to nine large job categories as well as in relation to the particular job proposed for him.

As a servicing agency the chief aim of the assessment board was to discriminate unfitness, and so, in this respect, its role could be compared to that of the conventional screening unit, were it not for two significant differences: first, the OSS board had to appraise the relative usefulness of men and women who fell, for the most part, in the middle and upper ranges of the distribution curve of general effectiveness or of one or another special ability, people who had already been selected because of the skill they had manifested in some field of activity; and, second, these men and women had to be assessed in respect to a number of additional personality qualifications — social relations, leadership, discretion (disposition and ability to withhold confidential information), and so forth. The usual screening devices, based on elementalistic principles, though efficient in distinguishing people who, because of one or another marked defect or handicap, are incapable of functioning effectively, are not so successful in discriminating degrees of effectiveness among those whose test scores fall above the usual level of acceptability. Furthermore, none of the conventional screening devices are designed to test social qualifications—good will, tact, teamwork, freedom from disturbing prejudices, freedom from annoying traits, and so forth.

Also, besides eliminating the candidates who did not measure up to the relatively high OSS standard of overall acceptability, the assessment board was expected to predict the performance level of each candidate in fulfilling a designated role in a designated area of operations (e.g., branch chief in Cairo, leader of native resistance groups in Burma, etc.) and, in addition, to write a personality sketch of the candidate, which would assist his superior officer, or supervisor, in dealing with him, or in deciding to which specific place, specific group, and specific function he should be assigned.

To accomplish all this it was id-

ered necessary to set up a program of assessment according to organismic (gestalt) principles, that is to say, to design a variety of tasks-in-situations which would test a man's effectiveness in performing functions of the same type and at the same integrative level, and under somewhat similar conditnons, as those he would be required to perform in the field. For OSS personnel this meant, among other things, that a candidate would have to be confronted by a succession of *varied novel* assignments—requiring quick planning, coordination of muscles with or without instruments and/or the coordination of ideas and words—many of which assignments would have to be executed in collaboration with other people. Make-believe, yet to a degree, realistic test situations conform to organismic principles in so far as they call for functional operations, or *proceedings*, of personality at a relatively high level of differentiation and integration, the level that must be sustained if one is to act appropriately and efficiently in everyday life, especially in a policy-making, administrative or executive capacity.

The assessment board also complied with organismic principles in so far as they attempted in each case to synthesize significant observations of behavior into partial formulations of the personality of the candidate, and from these formulations to deduce the behavior that could be expected of him in this or that situation in the theater. In other words, the task of assessment, as an intellectual proceeding, was a matter of diagnosis (inferring general tendencies and traits and their interrelations from a number of specific signs), and prognosis (on the basis of what is known about these tendencies, traits and relations, to predict future patterns of behavior) and finally, with the latter in mind, to decide on the suitability of the candidate. This differs from the elementalistic mode of appraisal which consists of a series of objective measurements of relatively simple processes or isolated traits, which leads to a final decision based on a list of scores weighted in connection with known correlations between scores and levels of performance (as determined in previous studies of similar groups). The elementalistic method does not call for a diagnosis of the personality.

The bulk of OSS assessment in the United States was done at Station S which was located on a hundred-acre farm, fifteen miles from Washington, and at Station W in Washington, D.C. The former had a three-and-a-half day program, the latter a one-day program of procedures. This paper is limited to an account of the conditions that prevailed and the techniques that were employed at Station S.

There a senior staff of about 7 psychologists, psychiatrists, and sociologists, and a junior staff of about 8 graduate students of psychology lived intimately with groups of approximately 18 candidates, who, in the course of their three and a half days in the Area, were entertained and tested and assessed in a variety of ways.

Regardless of military rank or civilian status, all candidates arrived at S in army fatigues and during the period of their stay hid their identities from each other by means of fictitious names and stories of their lives.

The final evaluation of each candidate and the judgment of his suitability for his proposed assignment came from a number of sources and procedures:

1. Casual observations. These were made constantly, since staff and candidates lived together in one house, ate in the same hall, relaxed in a common livingroom, and played games together

both in- and out-of-doors.

2. Standardized tests of abstract intelligence, of mechanical comprehension, etc. These were paper-and pencil tests of both group and individual type.

3. Projective tests, such as Sentence Completion, Thematic Apperception Test, Rapid Projection Test, the Rorschach, and so forth, used primarily to uncover repressed tendencies which might affect a candidate's motivation or emotional stability.

4. Personal history interviews with each candidate lasting at least an hour and a half, in preparation for which the interviewer examined a Personal History Form filled out by the candidate, his standard test scores, and his performance on the projective tests, and in the course of which the interviewer sought insight into the structure and dynamics of the candidate's personality and his motivational, emotional, and intellectual fitness for his assignment.

5. Group situational tests in which 4 to 7 candidates were faced with the task of solving a problem, either indoors or outdoors, in some cases with a group leader assigned and in other situations with no leader designated. Group situations of this sort were first introduced by the British W.O.S.B.'s (War Officer Selection Boards) in their testing of officer candidates.

In Leaderless Group Situations the development of group structure could be observed as well as the contribution of each man to it, his practical intelligence, his leadership, his energy and initiative, his social relations, and in outdoor situations, his physical ability. Typical leaderless group situations at S were:

a) Wall Test. A group of 6 men were brought to one of two parallel walls, 8 feet apart and 10 feet high, and told that their task was to get themselves and a "king-size bazooka" (actually a heavy 8-foot log) over the two walls and down on the far side without getting down between the walls.

b) Discussion Test. A group of six men, sitting around a table, were asked to discuss for half an hour the question, "What are the major postwar problems which the United States will face after the war?" and at the end of that time, to choose one of their group to present their conclusions to the staff.

In Assigned Leadership Problems, each member of the group was, in turn, assigned to lead his group in solving some problem, for example, crossing a mined road, searching an agent's room in enemy-occupied territory, and so forth, situational tests in which fact and fancy were fused.

6. Individual Situational Tests in which a candidate met a situation alone, such, for example, as:

a) Construction Test. The candidate was shown a glorified tinker-toy with which he was instructed to build a 5-foot cube with 7-foot diagonals in ten minutes with the aid of 2 helpers whom he was to direct. The "helpers" were members of the junior staff who soon turned the situation into a test of the candidate's frustration-tolerance by becoming increasingly lazy, recalcitrant, and insulting. No candidate ever completed this task, and there were those who became either markedly upset or enraged by the humiliations they suffered.

b) Stress Interview. The candidate was told to assume that he had just been caught going through a secret file in a government building in Washington, and, after ten minutes in which to think up an innocent and plausible explanation for his presence there, he reported to a basement room where, under the glare of a spotlight, he was given a grilling cross-examination on the details of his story in an attempt to confuse and disquiet him as much as possible.

7. An obstacle course to test the candidate's physical ability, especially his daring and agility.

8. Tests of observing and reporting, for example, the terrain at S, personal effects left in a room, a map, contents of a brief case, etc.

9. Tests of propaganda skills, for example, a script for a spot radio broadcast and a pamphlet to disturb Japanese workers in Manchuria.

10. Psychodrama. Situations tailor-made for each candidate to reveal his ability to cope with difficult social situations *vis-a-vis* another candidate.

11. Debate. Toward the end of the program a convivial evening party was held with hard liquor for both candidates and staff with a more or less informal discussion or debate by the candidates on a timely topic, such as, "What are we fighting for?" "What should the Allies do with Germany after the war?" etc. The party which lasted usually long past midnight often revealed aspects of the candidates' personalities not seen in soberer states.

12. Sociometric questionnaire given candidates on the last day which revealed their opinions of the colleagues, those whom they would choose as their leaders, those whom they would accept socially and those whom they would reject, those who were most dogmatic, most inconspicuous, most irritating, etc.

13. Judgment of others. As a test of their ability to size up others each candidate was asked to write a personality sketch for each of the five men in the group whom he had come to know best during his stay at S.

Material from all these sources, interpreted in the light of the picture of the man obtained in the personal history interview, and, in turn, correcting that picture, was utilized in writing the Evaluation Note on the candidate which constituted the main part of the S Final Report. In this note of six to seven hundred words an attempt was made to convey in nontechnical terms the main features of the character structure of the candidate.

A second part of the Final Report consisted of ratings on a six-point scale (very inferior; inferior; low average; high average; superior; very superior) of ten variables of personality, the first seven of which were considered important for all candidates, the last three of which were necessary for certain types of assignment.

These variables were:

1. *Motivation for Assignment*: the degree to which a man felt impelled to contribute to the war effort; his interest in the proposed assignment.
2. *Energy and Zest*: average level of purposive activity and initiative.
3. *Practical Intelligence*: skill in dealing with things, with people and with ideas.
4. *Emotional Stability*: goverance of emotions under stress; absence of anxiety, anger and depression; absence of neurotic symptoms.
5. *Social Relations*: ability to get along with all kinds of people; cooperativeness; absence of annoying traits.
6. *Leadership Ability*: ability to organize groups and evoke cooperation and respect.
7. *Security*: ability to keep secrets and maintain cover.
8. *Physical Ability*: agility, daring, ruggedness.
9. *Observation and Reporting*: ability to observe and remember significant facts; to evaluate information; to report accurately and concisely.
10. *Propaganda Skills*: ability to devise and execute subsersive measures for disintegrating the morale of the enemy.

In order the more adequately to observe candidates and to rate them on the variables of personality relevant for their assignments each class of 18 was divided into 3 subgroups of 6 men, to which a team of 2 senior and 1 junior staff members was assigned. It was the responsibility of this staff team to interview their 6 candidates and to rate them in various situations and tests, and to prepare the Final Report on each of them. This provided that the ratings of all personality variables of each candidate rated in a situation was the result of at least 3 independent ratings combined with final ratings in substaff conferences in which the performance of each candidate and its meaning were discussed at some length.

Since there were, in all, 35 procedures, and since each procedure provided an opportunity to rate 1 to 4 variables, there were anywhere from 2 to 10 scores for every candidate on each of the 10 variables by the end of the three and a half days of assessment. These ratings for each candidate were represented by colored pins placed on a huge scoreboard at one end of the staff room. For every rating above average a red pin was inserted, and for every rating below average a blue one, thus yielding an easily grasped picture of each man's performance in all procedures at S at the time his case was discussed by the entire staff at its final meeting on the afternoon and evening of the last day of a class.

At this staff meeting it was the custom to have the first part of the Evaluation Note on a candidate read by the senior staff member who had interviewed him and who was responsible for his overall evaluation, and to have the second part, which described the candidate's behavior in the situations at S, read by the junior staff member who had been responsible for this part of the note. Other staff members contributed their impressions of the candidate in question, and proposed changes of or additions to the note. When the final form of the Evaluation Note had been agreed to by the staff as a whole, the final ratings of the variables for this candidate were made and a decision reached as to his suitability for his proposed assignment. The Final Reports made up of these three parts—the job fitness rating, the ratings of the 10 personality variables, and the Evaluation Note—were in this way prepared for each candidate, typed, and already delivered in Washington before the candidates reported there the next morning.

About 5,500 men and women went through the Assessment Centers in this country. Of these, about 20 per cent were not recommended for overseas duty, and another 20 per cent were not sent overseas for other reasons. Of those who did go overseas, about 1,200 were followed up in the different theaters – ETO, METO, IBTO, and CTO—where they were appraised by those who had supervised their work, as well as by 3 or 4 of their associates.

The analysis of this appraisal has not been completed. At the present time a report can be made only on the first 300 cases. On this basis, 6 out of every 100 men passed by the assessment staff proved to be unsatisfactory in the theater for one reason or another. Some of these unsatisfactory persons had been given assignments in the field other than those for which they had been assessed. Some others had failed to measure up to the high OSS standard because they lacked the necessary degree of some very specialized skill, the determination of which did not fall within the province of the assessment program. Only one of the 300 cases failed because of a neuropsychiatric condition.

# DEVELOPMENTS IN OCCUPATIONAL CLASSIFICATION

BY CARROLL L. SHARTLE
OHIO STATE UNIVERSITY

IN TIME of war when millions of persons shift from one kind of occupational activity to another it is necessary to have information regarding such occupational assignments and to classify occupations in a systematic fashion for ready use. A knowledge of such occupations is necessary for a proper interpretation of test scores, personal data, work experience and other information for use in the classification and assignment of personnel in military and civilian activity.

Occupations can be classified according to a large number of variables singly or in combination. Roughly the variables are of three types:

A. Variables concerning the characteristics of individuals who are or have been employed in the occupations.
B. Variables concerning the characteristics considered important for individuals to possess who enter the occupations.
C. Variables which are primarily the characteristics of the jobs in the occupations rather than of individuals who have been employed in the occupations, are now employed or who may be selected later to fill such jobs.

Most occupational classification schemes are complex and are determined on the basis of several variables usually of the B and C types. Psychologists have been more frequently concerned with occupational classification problems involving A and B or combinations thereof, rather than with schemes involving C variables.

The war researches which have been published thus far have been more predominately of the B C types, although there has been a wealth of information collected which should result in the classification of occupations according to the characteristics possessed by workers employed in the occupations. A recent study by Harrell and Harrell [5] reported the Army General Classification Test (GCT) scores for 18,782 white enlisted men in the Army Air Forces Air Service Command. Tables were provided showing the distributions of scores according to previous civilian occupations. No attempt was made to arrange the occupations into a systematic classification structure, perhaps because only 74 occupations were covered and also because of the overlappings even though some differentiation was reported. It is hoped that other similar studies will be reported by the Armed Forces and also by civilian governmental agencies. While the results of a single test usually reveal considerable overlapping, such information is necessary to show the limitations as well as the differentiations obtained and to provide background for a realistic interpretation of test scores in terms of occupational groupings.

There was considerable activity during the war in classifying occupations according to various worker and job characteristics. The characteristics

were estimated and the occupations were classified for a specific purpose.

### OCCUPATIONAL LISTS

Three classifications of occupations were prepared which in published forms were occupational lists. There were Occupations Suitable for Women [7], Essential Occupations and Critical Occupations [11]. Each was designed for a specific purpose and although published as a single list several variables were considered in each case to determine proper listing. In the case of the Critical Occupations, for example, three variables were utilized:

1. Whether or not the occupation was in an essential war activity.
2. The length of time required to learn the occupation.
3. The degree of shortage of workers in the occupation.

The list of Critical Occupations with definitions was supplied to the Selective Service Boards and to the offices of the United States Employment Service as aids in the allocation of man power. The Essential Occupations list was an earlier and longer list of 2,000 occupations developed according to the first two criteria used for determining Critical Occupations. The titles given for the Essential Occupations were not defined but only occupations with a training time of six months or more were included. Needless to say both lists received wide use.

### MILITARY CLASSIFICATIONS

In order to facilitate the occupational classification and assignment of military personnel, both the Army and Navy made studies of military specialities.

The duties performed by Army Officers were studied by questionnaire method and about 800 occupational specialities were classified into 10 main groups with a new functional coding system to facilitate its use [8]. The Army enlisted men's occupational specialities were extended to a total of about 600 [6]. Although the code numbers assigned to titles did not follow a classification procedure the enlisted men's military specialities were grouped under 10 major fields of activity.

The Navy made studies of both officer and enlisted men's military activities and developed improved classification structures.

### DICTIONARY EXPANSION

The *Dictionary of Occupational Titles* [1] originally released in 1940 was expanded to include in its classification structure 21,653 occupations, an increase of 4,000. The principal activity in regard to the Dictionary (Parts I & II) was one of expansion with no basic change in the structure itself. This was also true of Dictionary (Part IV) The Entry Occupational Classification [2] which during the war was considerably enlarged and included several new features such as lists of Army, Navy and Civilian training courses which are related to the classification structure. Part IV is a classification of fields of work rather than specific occupations and in the post war period is receiving very great attention as a tool in vocational counseling.

### CLASSIFICATION IN INDUSTRY

A very great increase in occupational classification in industry developed during the war. First there was the necessity of having adequate classifications and definitions of jobs in order to justify Selective Service deferments. Second, because of wage controls job evaluation systems had an extensive growth.

In requesting occupational defer-

ments industry was required to show the necessity of the occupation and information regarding duties, replaceability, training time and other factors. Training time was perhaps the most important single criterion used. In addition to justifying each occupation deferment on a Selective Service form about 8,000 plants prepared Manning Tables [10] which showed the classification of each job in the plant including an estimate of training time and other factors pertinent to deferment.

Job evaluation systems became almost universally necessary in justifying proposed changes. The increase in the strength of labor unions was also a factor. The point system for job classification retained its lead as the most frequently used system in job evaluation programs.

### OCCUPATIONAL FAMILIES

Occupational families, which are constellations of occupations, developed considerably during the war. The method for establishing occupational families was published in 1940 [12]. Although some improvements were made in the technique the general method remained the same throughout the war.

The occupational families which were developed were of two types, military and civilian. Military occupational families were lists of civilian occupations which had proximity to a given military speciality. Such occupational families were used to aid in determining the best kind of civilian experience for the particular military speciality. The first such occupational family produced was for bombardier.

Occupational families were prepared for a number of civilian occupations to aid in the recruitment of war workers and the transfer of peace time workers to war production jobs. The occupational families thus produced were developed around a single occupation or on an industry basis. In the industrial occupational families a small number of related occupations was shown for each occupation in the industry. In a sense the industrial occupational families became abbreviated occupational families. Wide use was made of both the individual occupational family and the industrial occupational family, especially by the United States Employment Service and the War Manpower Commission. Individual plants likewise made application for such tables of related occupations.

The most recent development of occupational families was the preparation of aids for transferring military personnel in civilian jobs [3] [4]. Occupational families were developed for each military speciality. The number of related civilian occupations ranging from twenty-five to eighty. These occupational families have received wide use in the Army, Navy, United States Employment Service and by individual establishments. When used in connection with the *Dictionary of Occupational Titles* they present a wide variety of civilian occupations for each military speciality.

### SUMMARY OF DEVELOPMENT

In reviewing the developments of occupational classifications during the war it is apparent that the chief progress was made in expansion rather than in the development of new methods and techniques of classification. Before the war government agencies and the Armed Forces both had good foundations for enlarging occupational classifications and listings in order to meet their needs. It is unfortunate that more variations in technique were not tried out as no classification scheme is ever found completely satisfactory for the purpose for which it was developed. The

most promising classification methods for the future appear to be in the field of psychology in which there will be greater emphasis on the worker traits involved in classification rather than upon duties performed in the occupation.

BIBLIOGRAPHY


1. *Dictionary of Occupational Titles*, Parts I, II, and Supplement, 1939 and 1944. Washington, D. C.: U. S. Government Printing Office.
2. *Dictionary of Occupational Titles*, Part IV, Entry Occupational Classification (Revised edition, 1944). Washington, D. C.: U. S. Government Printing Office.
3. Division of Occupational Analysis, War Manpower Commission, *Special Aids for Placing Military Personnel in Civilian Jobs* (Enlisted Army Personnel), 1944. Washington, D. C.: U. S. Government Printing Office.
4. Division of Occupational Analysis, War Manpower Commission, *Special Aids for Placing Navy Personnel in Civilian Jobs*, 1943. Washington, D. C.: U. S. Government Printing Office.
5. Harrell, Thomas W. and Harrell, Margret S. "Army General Classification Test Scores for Civilian Occupations," *Educational and Psychological Measurements*, Autumn, 1945, 5 (3): 229-39.
6. *Military Occupational Classification of Enlisted Personnel (TM 12-427)*, 1944, Washington, D. C.: U. S. Government Printing Office.
7. Occupational Analysis Section, U. S. Employment Service, *Occupations Suitable for Women*, 1942.
8. *Officer Classification Commissioned and Warrant (TM 12-406)*, 1943. Washington, D. C.: U. S. Government Printing Office.
9. Shartle, Carroll L., Dvorak, Beatrice J., Heinz, Carl L. and others, "Ten Years of Occupational Research," *Occupations*, 1944, 22 (7).
10. Shartle, Carroll L., "The Manning Table —A Personnel Blue Print," *Occupations*, 1943, 22 (2): 20-23.
11. Shartle, Carroll L., *Occupational Information—Its Development and Application*. New York: Prentice-Hall, 1946. Pp. 96-101.
12. Stead, William H., Shartle, Carroll L. and Associates, *Occupational Counseling Techniques*. New York: American Book Company, 1940. Pp. 175-206.

# INDUSTRIAL IMPLICATIONS OF WARTIME DEVELOPMENTS IN PSYCHOLOGY

BY MORRIS S. VITELES
UNIVERSITY OF PENNSYLVANIA
CHAIRMAN, COMMITTEE ON SELECTION AND TRAINING OF AIRCRAFT PILOTS, NATIONAL RESEARCH COUNCIL

IN A LARGE SENSE, the beginnings of industrial psychology are to be found in the application of psychology for military purposes in World War I. Munsterberg's formulation of a program of industrial psychology appeared shortly before the opening of that war. The impulse towards the application of psychology in industry gained strength from the successful use of psychological techniques in the selection and classification of Army personnel by almost every nation engaged in the war. Especially in the United States, the large-scale use of the well-known Army group intelligence tests and the development of tests for the assignment of military personnel to specialized military tasks stimulated the rise of a program of industrial psychology directed, in particular, to the selection and classification of industrial personnel.

The outcomes of research and practice in military psychology during World War II may well become the starting point for a new extension of psychological services in industry. The effect of World War II has already been felt in the form of an increase in the number of industrial firms making use of psychological methods for the selection and classification of personnel. Evidence of such an increase is found in a survey recently conducted by the author, in collaboration with A. S. Thompson, on the use of psychological tests by industries in the Philadelphia area [19]. This survey was made by means of a questionnaire dealing with the nature, size, and administration of the selection test program sent to 93 firms representing those holding membership in the Industrial Relations Association of Philadelphia. Returns received from 64 firms, representing 68.8 per cent of the total number contacted, showed that between the end of 1941 and February 1945, the number of companies using psychological tests had increased from 18 to 29. This increase (Table I) during a period of somewhat slightly more than three years was equal to that occurring during the preceding eight-year period (1934-1941), and greater than that of the earlier eight-year period (1925-1933).

TABLE I

NUMBER OF COMPANIES IN PHILADELPHIA USING PSYCHOLOGICAL TESTS FOR SELECTION AND CLASSIFICATION OF EMPLOYEES

| Year | No. | Year | No. |
|---|---|---|---|
| 1925 | 3 | 1936 | 9 |
| 1926 | 3 | 1937 | 10 |
| 1927 | 4 | 1938 | 11 |
| 1928 | 6 | 1939 | 12 |
| 1929 | 6 | 1940 | 15 |
| 1930 | 6 | 1941 | 18 |
| 1931 | 6 | 1942 | 23 |
| 1932 | 6 | 1943 | 25 |
| 1933 | 6 | 1944 | 28 |
| 1934 | 7 | 1945 | 29 |
| 1935 | 9 | | |

It is difficult to estimate the extent to which the situation in Philadelphia is

representative of the country at large. However, a survey reported by D. G. Marquis [5] in 1944 on the mobilization of psychologists for the war effort showed that there were 245 "qualified" psychologists, i.e. either members of professional associations or those qualified for membership in such associations, employed by business, industry, and consulting organizations. To this must be added a much larger number of less well trained individuals employed by many industrial and commercial firms throughout the country in applying psychology in the selection, training and maintenance of personnel. Principles, methods and materials used by the military forces during World War II can be employed to implement the work of this group and to broaden the basis for increased participation by psychologists in dealing with problems of employee adjustment, production, and morale in business and industry.

## SELECTION AND CLASSIFICATION OF PERSONNEL

Illustrations of the implications to industry of wartime developments in psychology are to be found in the area of employee selection and classification. For example, even in plants using psychological methods, it is commonly the opinion of management that a testing program covering the entire personnel of the plant is unwieldy, excessively costly, and perhaps generally impossible of attainment. Even in plants where psychological tests have been used over a period of years, the program has generally been limited to a relatively small number of jobs in which hiring or adjustment problems are particularly pressing. Perhaps one lesson that industry can learn from the wartime applications of psychology in selection and classification is that a large-scale, all-inclusive program is feasible within reasonable limitations of space and personnel and, when properly planned and validated, can produce a return proportionate to the cost of its administration.

Work done by both civilian and military groups in the interest of the war effort illustrated the practicability of the large-scale application of psychological methods for the assignment of personnel to specialized tasks. Approximately four years ago, the War Training Service of the Civil Aeronautics Administration was asked to provide for the Army numbers of trained pilots to fly small planes; to act as pilot instructors; to prepare for glider operation, etc. The National Research Council Committee on Selection and Training of Aircraft Pilots, which had previously carried on considerable research on pilot selection with funds made available by the Civil Aeronautics Administration, undertook to provide uniform administration and scoring on a nation-wide basis of tests for screening candidates for the proposed training program. This was done through the establishment of a National Testing Service, centralized at the University of Rochester, under the direction of J. W. Dunlap, which, in the course of approximately nine months was responsible for the examination, by 609 co-operating examiners, of 67,067 candidates for pilot training, reporting to 571 co-ordinators scattered throughout the United States. The administrative cost of this program, involving the application of four pencil-and-paper tests, amounted to approximately $75,000. In other words, it was found possible to apply scientific techniques instead of "hunches" and "guesses" in a program whose magnitude would have frightened the average psychologist in the field only a few years ago at the extremely low cost of little more than one dollar per candidate for training [7].

Even this program represents a small operation in comparison with that of the Army Air Forces selection and classification program, in which, under the direction of Colonel J. C. Flanagan, "assembly line" methods have been employed in running hundreds of thousands of cadets through a series of psychomotor tests and applying the results with outstanding success in determining whether a man could most competently serve as a pilot, bombardier, or navigator [11].

It is of interest to note that an operation of this magnitude in the classification of personnel was made possible through the use of test batteries and simultaneous differential classification for a series of tasks. The usual practice in industry, in the past, has been to use a test battery for a single job. Developments in the formulation of "job families" and actual experience of the Army Air Forces and other organizations point to the possibility of using a single battery to determine for which of a number of jobs a man is suited [10]. In the Army Air Forces classification program, the number of occupations has been limited to three. Simultaneous testing for only three jobs would greatly facilitate the industrial testing program, but there is no reason why, with the use of proper validating techniques, the number cannot be extended to include ten, twelve, or even more jobs. It can be readily seen that such a program will reduce the cost of testing and make possible the large-scale operation, covering the entire company, which has just been described as feasible and practical for industry. In passing, it is of interest to note that the Occupational Analysis Section, War Manpower Commission, has already undertaken research in this area. One outcome reported by that agency is a standard aptitude test battery involving 24 tests requiring approximately three hours, intended for use in counseling veterans, which can be scored for approximately 90 specific occupations [9].

Developments have also occurred in the analysis of specialized techniques for selection and classification which have important implications for industry. Studies made in industry before World War II have indicated the potential value of biographical items, available from application blanks, in the selection and placement of personnel. However, industry has largely confined the use of weighted application or biographical item scores to the selection of sales personnel. Evidence available from wartime research, leading, for example, to the use of such items in the selection of aviation cadets [8], points to the advantages of extended research in this area, which can often provide a tool less costly and less cumbersome than psychological tests and considerably more effective in its return.

The interview represents another specialized area of investigation conducted under military auspices which will provide data and principles useful in the application of psychology in industrial selection and classification. Representative of such investigations is an experiment, by the National Research Council Committee on Selection and Training of Aircraft Pilots at Harvard University, Ohio State University, Purdue University, and the University of Michigan, in which a Board of Interviewers conducted a relatively standardized interview involving the use of a carefully devised rating scale [2]. Agreement among interviewers in predicting success or failure proved to be quite high, mean correlations among independent interviewers in rating "Fitness for Flight Training" being close to .90. The interview fared relatively well in predicting certain objective cri-

teria of competence in flying. However, while it met the routine tests of scientific acceptability, the interview failed of practical justification on a most basic and critical point. Because of the high cost of the interviewing, the interview, to be acceptable as having practical usefulness in selecting pilots, must add significantly to the predictive efficiency that is obtainable by the application of pencil-and-paper tests to large groups. The interview failed to do this. Using the paper-and-pencil tests, to which reference has been made, predictions can be made for 500 men with a total time expenditure of about 2 to 10 man-hours. Adding an individual interview to this prediction did not materially increase the efficiency of prediction, although it added at least 750 man-hours to the time expended in assuring competency for flying. Such findings have great practical and economic significance with respect to the use of the interview and research on the interview in industry.

Wartime research has also been responsible for the production of a number of new tests which may find useful applications in industry. The number of the latter may actually be fewer than anticipated, since much of the productive work that has been done in the military service has involved the use of instruments which were already available before the war started. Nevertheless, there will be old tests that have been improved and refined and new tests which can be employed in industry. However, the experience of the military services will only serve to underline again the fact that tests and test standards cannot be arbitrarily transferred from one situation to another.

If anything, the experience shows that a repetition of the mistake made after World War I in moving bodily material and testing procedures developed by the Services into industrial plants without further evaluation will only reproduce the same failures that grew out of this procedure after the last war. In wartime applications of psychology, the psychologist has become increasingly aware of the need for validation and cross-validation in any situation other than that for which the test was originally devised. The quality of industrial research can be improved through the consideration of such experience in the development of programs for the application of psychology in the selection and classification of industrial personnel.

### IMPROVEMENTS IN CRITERIA

In the military situation, as in industry, great difficulty has been experienced in obtaining trustworthy estimates or measures of achievement during training and of performance subsequent to training as a means of validating predictors. For example, at first glance, the problem of measuring accuracy on the part of anti-aircraft gunners in hitting a sleeve towed by a slow moving plane may seem to be a relatively simple one. However, experiments conducted by the author showed this to be a most complex situation which was further complicated by the question as to whether accuracy in hitting a sleeve from the firing line has any relationship to performance in shooting down fast planes engaged in offensive or defensive tactics [17].

Similar difficulties were experienced in many other areas. Nevertheless, great strides have been made in the improvement of criteria. For instance, in the selection of pilots, instructors' judgments as to the progress of the student, inspector scores given at the end of flight training, etc., have been found to have as little reliability as the foreman's gross estimate of employee performance. However, there is evidence

from studies conducted by H. A. Edgerton and R. Y. Walker [3] leading to the construction of the *Ohio State Flight Inventory*, that the general character and reliability of direct observations of flight performance by instructors and inspectors can be improved by providing specially designed forms calling for ratings on specific aspects of flight performance. In aviation research such subjective criteria have been supplemented by photographic recording of instrument readings and of control movement which, when subjected to analysis through the use of methods developed at the University of Pennsylvania by the author and A. S. Thompson, give clearly objective pictures of pilot performance and plane response [18]. The measurement of flight performance has been revolutionized through the application of standard flights [20] which insure comparable situations in obtaining information on the student's progress in learning to fly and on the quality of his flight performance.

In other areas, standard work samples on operations actually performed have been developed and used to advantage both in the validation of tests and as items in measuring progress in training. There has also been considerable progress in the analysis of varied rating and nominating techniques, including, as in studies by Captain J. G. Jenkins, applications to the combat situation which have great significance in terms both of methodology and practicability. Industry in the postwar world can profit greatly from such advances in the evaluation of performance which have characterized the application of psychology in the war effort.

## PSYCHOLOGY AND EMPLOYEE TRAINING

The necessity for rapidly training large numbers for highly specialized military tasks focussed attention during World War II upon the problems of learning and teaching in the military situation. As a result, the work of the psychologist in the War effort will provide industry with numerous examples, in the way of training programs and training materials, showing how the fundamental principles of learning can be applied to shorten the training program and improve its outcomes. More important still, the outcomes of research in the military situation can help to set the pattern for the experimental determination of the best conditions for speeding industrial training and achieving optimal levels in terms of quality and quantity of output. There is, for example, evidence from research conducted under the direction of D. B. Lindsley, that a wrong method of instruction may actually retard learning which immediately takes an upward trend following an evaluation in instruction method. There are data from investigations by G. K. Bennett and his group showing how a change from uniform to varied instruction affects performance in learning radio code. Studies by the author and his associates on a National Defense Research Committee project [15] indicate the relative ineffectiveness of using synthetic devices to teach range estimation as compared with instruction in range estimation on the firing line. Such results and many others, representing particularly the outcome of research carried on through the Applied Psychology Panel of the National Defense Research Committee, described elsewhere in this symposium, indicate the benefits that can be achieved by industry from the extension of training research in the industrial situation.

Industry can also learn much from the attention centered by the military psychologist upon the teaching process. It is true that wartime industry profited greatly from the widespread dissemina-

tion, through the War Manpower Commission Job Instructor Training programs of basic principles of good instruction and other information on how people learn. However, while the Job Instructor Training course has been widely accepted, industry has not generally taken the next step of obtaining exact information on methods actually used by different instructors in training their men and on individual differences among instructors which are pertinent to the training situation. How this can be done is well illustrated in research conducted by E. L. Kelly, Purdue University, for the National Research Council Committee on Selection and Training of Aircraft Pilots. In this, a short-wave transmitter and electrical interphones were installed in a plane and instruction as given by the instructor recorded on the ground. In passing, it is interesting to note that this method of transmission was later replaced, largely through the co-operation of E. L. Kelly, Purdue University, and D. R. Brimhall, Civil Aeronautics Administration, with the Armour Research Foundation, by an airborne model magnetic wire recorder, weighing only twelve pounds, which does not necessitate short-wave transmission and all its attendant difficulties.

One important outcome of such research was an opportunity to examine the variations among instructors with respect to the terminology used in training pilots [4]. For instance, an analysis of ten hours of instruction by each of four instructors revealed a total of 500 technical or specialized terms, many of which were used only by an individual instructor. In addition, it was found that much instruction was given in the air which could better be given on the ground and that, in general, pilot training suffered from inadequate methods of presentation by instructors who did not apply the fundamentals of good teaching procedures. The outcome was a series of manuals containing material to be covered on the ground and in the air, including standard "patter" for instruction and a standard course of training methods applying to the instruction of pilots [21].

Such research, which has been duplicated in connection with other military tasks, has resulted in the production of manuals and methods which can be useful to industry. "How to Teach Gunnery" and "How to Teach Engineering" prepared at the University of Pennsylvania for the U. S. Navy [16, 14] are illustrations of such materials. It may well be, for example, that "key points" on jobs are not made clear because different instructors use different terminology in explaining how the job is done, and some of the terms are highly technical or highly individual. The investigation of such differences and of individual variations in teaching methods can produce for industry, as it has for the military services, fruitful returns in the way of reducing training time and in improving the quality of performance.

## MEASURING AND INFLUENCING EMPLOYEE ATTITUDES

Of particular significance to industry is the progress which has been made in the applications of social psychology in the interest of the war effort. The procedures of this specialized branch of psychology have been used in investigating the attitudes of military personnel and in testing the adequacy of a program designed to influence the states of mind and the reactions of individuals and of groups. For example, with techniques developed in peacetime research, the Morale Services Division of the War Department has measured what the soldier wants, how he reacts to his Army experiences, what gripes him most and why, and what he would like to see done

about it. These studies are made because research has demonstrated that soldier attitudes underlying morale can be improved markedly by meeting his reasonable desires and needs in a systematic fashion [6].

Such surveys have been supplemented by a systematic program of experimental studies designed to measure the effects of particular programs formulated to influence attitudes and opinions. In the conduct of such research, which has extended over many areas, improved instruments and methods of measurement have been developed. Among these is a new theory of scale construction, growing out of practical needs, which gives promise of wide utility [13]. Even before the war, industry made considerable progress in the use of such procedures for the measurement of employee attitudes. It seems quite likely that what has been learned during the war can be employed to advantage in improving and extending this aspect of personnel work in industry.

## PSYCHOLOGICAL ASPECTS OF MACHINE DESIGN

There is one other respect in which industry can profit from the wartime applications of psychology. Personnel workers have largely thought of their function as one providing a competent man to operate tools and machines. In a large sense, training may be viewed as a method of adapting the man to the machine [1]. However, there is growing evidence that the need for speed and efficiency in industry and the conditions for the adjustment of the worker can better be satisfied if more attention is given to the adaptation of the machine to man.

This is a principle well recognized in industrial psychology, and yet constantly neglected, largely because in an era of technology, engineers, concerned with the development of the machine, and psychologists, concerned with the individual who is to operate the machine, have worked in almost complete isolation from one another. Not only are provisions for direct contact generally absent, but conditions have been such as almost to ensure against any danger of a possible short-circuiting of ideas between the two groups. Disregard of basic principles of perception, of basic patterns of motor co-ordination, of human limitations in the integration of complex responses, has led to the production, for both civilian and military usage, of mechanical monstrosities which could have been avoided if there had been co-operation between psychologist and engineer in the process of designing the machine.

While the application of psychological techniques for identifying selected individuals capable of operating such equipment has alleviated the situation, there nevertheless remains the basic problem of designing apparatus in advance so that its operation is less difficult and requires less in the way of selected personnel. In the military services, as a result of research of the type directed by D. B. Lindsley, Brown University; C. W. Bray, Princeton University; S. W. Fernberger, University of Pennsylvania; W. J. Brogden, University of Wisconsin; W. E. Kappauf, Princeton University; Lieutenant Colonel A. W. Melton, Army Air Forces School of Aviation Medicine, and others, there has developed increasing dependence upon the psychologist, working in close co-operation with engineers, for advice on the suitability of a proposed weapon, vehicle, detector, and other varied equipment in the design stage. The problem of machines not suited to the men who are to operate them exists perhaps on a larger scale in industry than in the armed services. It is such as to call for

close collaboration between the psychologist and the engineer not only in the initial design of the machine, but also for the determination, as an essential feature of good personnel management, of those speeds and rhythms and other conditions which, in the operation of the machines, are optimal from the viewpoint of the welfare as well as the productivity of the worker.

REFERENCES


1. Burt, C. "Psychology in the War: The Military Work of American and German Psychologists," *Occupational Psychology*, 1942, 16: 108.
2. Dunlap, J. W. and Wantman, M. J. *An Investigation of the Interview as a Technique for Selecting Aircraft Pilots.* Washington, D. C.: CAA Airman Development Division, Report No. 33, August 1944.
3. Edgerton, H. A. and Walker, R. Y. *History and Development of the Ohio State Flight Inventory, Part I: Early Versions and Basic Research,* Washington, D.C.: CAA Division of Research, Report No. 47, July, 1945; *Part III: Recent Versions and Current Applications.* Washington, D. C.: CAA Division of Research, Report No. 51, November 1945.
4. Kelly, E. L. *The Flight Instructor's Vocabulary.* Washington, D. C.: CAA Division of Research, Report No. 22, October 1943. Pp. $x + 57$.
5. Marquis, D. G. "The Mobilization of Psychologists for War Service," *Psychological Bulletin*, 1944, 41: 469-73.
6. ———. "Social Psychologists in National War Agencies," *Psychological Bulletin*, 1944, 41: 115-26.
7. National Research Council, Committee on Selection and Training of Aircraft Pilots. *The CAA National Testing Service (Summary of Test Results and Comparisons with Success in Flight Training).* Washington, D. C.: CAA Division of Research, Report No. 29, November 1944.
8. From the files of the National Research Council on Selection and Training of Aircraft Pilots.
9. *Report to the Technical Board, Highlights of Our Work, Analysis and Manning Tables.* War Manpower Commission, June 1943-February 1944.
10. Shartle, Carroll L., Dvorak, Beatrice J., Heinz, Carl A., Osborn, Martha L., Keller, Kathleen, Giese, Willis E., Pimm, Walter B., Nichols, William O., Lewis, Leon, and others. "Ten Years of Occupational Research," *Occupations*, 1945, 22, 409-14.
11. Staff of the Psychological Branch, Office of the Air Surgeon, Headquarters Army Air Forces, "The Aviation Psychology Program of the Army Air Forces," *Psychological Bulletin*, 1943, 40: 759-69.
12. Staffs, Psychological Research Unit No. 2 and Department of Psychology Research Section, School of Aviation Medicine, "Research Program on Psychomotor Tests in the Army Air Forces," *Psychological Bulletin*, 1944, 41: 307-21.
13. Stouffer, S. A. *Social Science and the Soldier in American Society in Wartime.* (W. F. Ogburn, Ed.) Chicago, Ill.: University of Chicago Press, 1943.
14. Viteles, M. S. How to Teach Engineering. A Short Course in Effective Teaching Methods for Engineering Instructors.
15. Viteles, M. S. and others. Evaluation of Methods of Training in Estimating a Fixed Opening Range. Part I: The Synthetic Trainer, Device 5C-4. Part II: Training on the Firing Line. Part III: Retention of Range Estimation Ability.
16. ———. How to Teach Gunnery. A Short Course in Effective Teaching Methods for Gunnery Instructors.
17. Viteles, M. S., Gorsuch, J. H. and Wickens, D. D. History and Final Report of Project N-105 Applied Psychology Panel (Final Report Under Contract OEMsr-700.)
18. Viteles, M. S. and Thompson, A. S. *An Analysis of Photographic Records of Aircraft Pilot Performance.* Washington, D. C.: CAA Division of Research, Report No. 31, July 1944. (Restricted.)
19. ———. "The Use of Psychological Tests by Industries in the Philadelphia Area," *School and College Placement*, 1945, 6: 1-4.
20. ———. *The Use of Standard Flights and Motion Photography in the Analysis of Aircraft Pilot Performance.* Washington, D. C.: CAA Division of Research, Report No. 15, May 1943.

# VALIDITY FOR WHAT?[1]

BY JOHN G. JENKINS
NAVY DEPARTMENT
WASHINGTON, D.C.

THE EVENTS of World War I taught American psychologists the *necessity* of validation. The experiences of the next two decades taught them much about the *technique* of validation. It remained for World War II to drive home to psychologists at large the necessity of devoting much time and thought to the *basis* for validation.

The necessity of validating predictor batteries has so long been accepted in this country that our younger colleagues often express surprise on learning that validation was ever regarded as anything but mandatory. Yet a review of German test literature of the years 1920-1940 will show that the German tradition has been to spend much time on test construction and then to ignore the problem of checking to learn whether the predictors do in fact predict. Enough has been learned about testing in the Luftwaffe and in the Reichswehr to make us sure that face validity, rather than measured validity, was regarded as adequate by the Germans even in war time. By way of contrast, tests developed and published without validation have never had any considerable acceptance in the United States.

[1] This article is written in response to the Editor's invitation to comment on significant developments in psychology in World War II. Instead of attempting to catalog the advances made by the 90 psychologists commissioned in naval aviation, the writer prefers to discuss a topic to the development of which so many members of this group made signal contributions.

Having accepted the obligation to validate, American psychologists made great strides, during the years 1920-1940, in learning *how* to validate. The advanced graduate student of 1940 had at his fingertips a variety of techniques which were completely or partially unknown to the expert of 1917. Item analysis was a commonplace. Factor analysis was available to many. Correlational measures of various sorts were familiarly used in every quarter. Textbooks provided correctional formulae by which the most pedestrian student could take care of truncation of range, attenuation, and other shortcomings of the standardizing sample. Techniques for combining predictors in the interest of maximal predictive efficiency were available for everyday use. Given a criterion, the psychologist of 1940 was prepared to do a very accurate job of predicting it.

It required the events of World War II to show many psychologists that the phrase "given a criterion" lies at the source of much inadequate prediction. Prior to the war—if the publications in our journals are any indication — psychologists in general tended to accept the tacit assumption that criteria were either given of God or just to be found lying about. They had earlier appeared to believe that little time need be devoted to what they were trying to predict, provided only that enough time were spent on the accurate combination of predic-

tor items, selected in the light of whatever criterion expediency might provide. These assumptions proved to be inadequate, when organized psychology launched itself upon the task of helping to select men for a war which the country appeared, for a considerable time, to be in serious danger of losing.

Any reader who feels that the above account is a caricature of the facts has only to turn to the published literature of 1920-1940. He will find many hundreds of articles dealing with the construction of predictors. When he comes to look for systematic treatments of the criterion-to-be-predicted, however, he will find the index strangely silent. During the 20 years immediately preceding World War II, the literature contains only a handful of titles dealing with the criterion and thoughtful enough to withstand critical scrutiny.[2]

The absence of systematic articles does not, of course, mean that psychologists here and there had not been thinking and even writing about problems encountered in securing adequate criteria. The point is rather that such writing was commonly buried away in articles dealing with specific predictive tests or batteries. The criterion was apparently not recognized by psychologists at large as a basic problem, warranting much careful thought and much insightful investigation on its own account. The novice of 1940, searching through many textbooks and much journal literature would have been led to conclude that expediency dictated the choice of criteria and that the convenient availability of a criterion was more important than its adequacy.

Psychologists working in Naval Aviation in World War II soon found that their criterion was not "given." Asked to help in forwarding to the combat areas as many "good" pilots as possible, they began to inspect and to discuss the whole problem of obtaining adequate criteria. By the end of the war, they had made certain advances in establishing criteria more nearly adequate to the demands of the sponsor.[3] The greatest advance, in the opinion of the writer, was however in the thinking and discussions about criteria which led this group toward the formulation of certain principles regarding the selection and development of criteria. This paper attempts to formulate briefly certain major outcomes of these discussions.

The basic postulate of adequate validational procedures may be stated as follows:

> Satisfaction of the needs of the sponsor can be expected only if the criterion employed is both reliable and valid in terms of the task set by the sponsor.

Reliability and validity of *predictors* is an old familiar topic to all readers of this journal. Applying the concepts of reliability and validity to the *criterion* is believed to be novel enough to warrant the discussion given below.

A somewhat more detailed statement of the basic postulate may be given in this form:

I. Criteria may be inadequately *reliable* because of:
   A. Instability inherent in the criterion

[2] The following articles are among those which stamp themselves as important: R. M. Bellows, Procedures for Evaluating Vocational Criteria," *Journal of Applied Psychology*, 1941, 25: 499-513; R. Hoppock, ed., "Criteria of Vocational Success—A Symposium," *Occupations*, 1936, 14: 917-75; P. Horst, *ed.*, "Prediction of Personal Adjustment," *Social Science Research Council, Bulletin* 48, 1941: 20-36.

[3] The term "sponsor" will be used throughout this paper to denote the person or agency originally requesting the establishment of selection procedures for a given task. Typical sponsors in wartime would be a training command, a bureau, or some cognizant official. In peacetime, sponsors include administrative officials in industry, and in the federal government. Ordinarily the sponsor lacks technical training and will express his needs for a selective program in nontechnical terminology.

performance itself.

B. Disagreement between allegedly competent judges on reviewing the same or successive samples of the same performance.

C. Administrative procedures leading to the making of critical entries by clerical personnel who did not observe the performance or who were inadequate to judge it.

D. Failure to secure a sufficient number of samples of performance or to observe the individual sample long enough to make a stable judgment.

II. Criteria may not be adequately *valid* to the task set by the sponsor because of:

A. Failure of the criterion-measure to comprise a large and significant part of the total field performance desired.

B. Failure of performance in training to forecast later field performance.

C. Intercurrent changes in training-performance due to

1. Advice given by the psychologist following his job-analysis or other observations.
2. Administrative or other extrinsic pressures.
3. Changes in the field performance itself.

These statements may be reviewed and illustrated in some detail. Most of the illustrative material will be drawn from Naval Aviation, since this group, by invitation, is represented in this article. At the same time, the writer is aware, through repeated discussions with professional colleagues who have served in other phases of the war effort, that much illustrative material can be presented from outside the Navy sources as the occasion permits.

## RELIABILITY OF CRITERION DATA

A. *Instability inherent in the criterion performance.*—Various predictive efforts in World War II have failed because the very performance that was to be predicted has proved to be inherently unstable. Number of hits scored by an aircraft gunner on a towed sleeve is a criterion possessing a delightful degree of objectivity. All studies of this criterion known to the writer, however, find the individual scores so low in reliability as to be without value in placing individuals in any sort of contiuum. Certain physiological measures encountered in Naval Aviation give high split-half reliabilities and very low test-retest reliabilities. Raymond Franzen, in comments made at professional meetings, has spoken of such measures as describing a biological instant in the life of an individual rather than the individual. It remains to be seen whether certain "personality" traits do not belong in this class. Certainly some of the research literature suggests a degree of inconsistency in these types of behavior which bodes ill for efficiency of prediction. It would not be surprising to find that here, as in the performances named above, we are dealing with criterion-reliabilities so low as to preclude the possibility of validities high enough for practical usefulness.

B. *Disagreement between judges.* — Student pilots pass or fail in aviation training on the basis of judgments by instructors and/or check pilots. In one good sized sample of Navy cadets, it was found that immediately successive check-flights, given by two experienced instructors, yielded a reliability of grading of approximately .00. (Actually the coefficient of correlation was slightly negative!) It is apparent that successive gradings cannot be predicted at this level of reliability. Research in categories of aviation training outside the Navy is known to have given similar results. In this situation, one either accepts the word of some one instructor and uses this as a criterion, knowing that other instructors would not agree; or one attempts to work with instructors to obtain improved agreement. It is to be hoped that a later article, by the

psychologist responsible,[4] will indicate what steps were taken in the Navy along the latter lines.

C. *Entries by clerical personnel.* — Navy instructors were required to fill out a detailed form, giving critical comments on various phases of each flight completed by each student. In actual practice, it was found that certain instructors (who were often conscientious in the actual instruction of their students) gave only a general grade, leaving the detailed grading to a clerk. The clerk, who had not seen the flight and could not have judged it acceptably if he had seen it, supplied the detailed grades in a fashion which gave an overall distribution of conventional form. The writer saw a close parallel in prewar days, when certain civilian instructors saved all their detailed forms to be filled in on a weekend. Thus, on a given Saturday, a civilian instructor might find himself faced with filling in a dozen details on a half dozen flights for each of ten or more students. The probability that his grade for a given detail for a given flight would agree with the grade that would have been given by another instructor for the same detail is presumably of chance order. Yet, in both of the above instances, the psychologist initially attempted to predict detailed entries in regard to coordination, judgment, observance of traffic rules, attitude toward flying, and the like. In view of the fact that the original entries were made as they were, one is tempted to describe them in terms of the concept of "Test-No-Test Reliability." The concept will be familiar to anyone who has reviewed thousands of physical examination forms only to find body temperature invariably set down as "98.6."

D. *Inadequate sample of performance.*—Early in the war, psychologists were invited to predict "failure" in the low-pressure chamber. They were confronted with the observable fact that certain men, when brought to simulated altitudes of 15,000 feet and above, either fainted or asked for discontinuance of the run. Research was to show that such "failure" was highly unreliable.[5] To obtain a dependable sample of those who could not tolerate altitude—among the young subjects in question—it was necessary to make repeated runs. Ultimately one ended up with a population consisting of those who failed consistently in repeated runs. Unfortunately, from a statistical point of view at least, this proved to be a minute portion of the sample. There is real reason to believe that failure in flight training, if subjected to the same procedure, could be made to yield a final (and quite possibly minute) population of "real" failures. The fact remains that, while flight failure is stable enough to predict, no battery has yet been found which significantly predicts one-time "failure" in the low pressure chamber.

## VALIDITY OF CRITERION DATA

A. *Criterion only part of field performance.*—Viteles tells the tale of how he set up a battery to predict the manual operations of a substation operator, only to find that the operator would be discharged if he failed to keep his station acceptably clean. The tale finds many parallels in the Armed Services. There is always the danger that the investigator may accept some convenient measure (especially if it be objective

[4] This study and the one reported later on the correlation between stages of training were conducted under the direction of Commander E. L. Kelly. He will presently publish accounts of these and related researches in aviation training.

[5] This series of investigations was under the direction of Commander R. H. Peckham. No unclassified account of it is available at this time.

and quantifiable) only to find ultimately that the performance which produces this measure is merely a part, and perhaps an unimportant part, of the total field performance desired by the sponsor. For example, the skilled aspects of piloting can be objectively recorded. Pilots could be selected on the basis of such criteria. To make such a selection, however, would be to ignore the importance of judgment and of emotional adjustment in piloting. We might, in other words, pick pilots who would do the right thing at the wrong time and in a state of emotional collapse—a selection which the sponsor would soon indignantly reject. Experience shows that the researcher must devote much time and thought to determining that the performance he selects as a source of criterion data is an adequate representation of the total field performance desired by the sponsor.

*B. Lack of relation between training and field performance.* — Performance in training has been a common criterion, chiefly because of its availability. As far as this writer knows, it has been the primary criterion for the validation of pilot-selection tests the world over. Granted that pilots who fail training cannot ever get into combat, we have no assurance that those who do well in training will do well in combat. There is at least a theoretical possibility that there may be a negative relation between training-performance and combat-performance. For example, by selecting for ability to pass training, we may be selecting conformers and rejecting the less conventional males, who might be our best bets in lethal air-assignments. Procedures to obtain data on this question have long been under way. In the meantime it seems to be incumbent on all of us to determine empirically whether, by using outcome of training as a criterion, we may be depriving our sponsors of able field-performers. This would appear to be a problem in the industrial and professional fields as well as in military occupations. It may be added, parenthetically, that surveys of this nature inevitably raise interesting questions as to the validity of the training procedures themselves.

C. *Intercurrent changes in the criterion.*—This may be regarded as involving either the validity of the criterion or its reliability, according to the definition given by the reader. It is here placed under validity simply because these changes imply that the performance that one predicts is inadequate to the task set by the sponsor as the sponsor currently conceives this task. Sometimes this comes about through the efforts of the psychologist himself. For example, the psychologists who were invited to work with the selection of aircraft gunners became convinced that the training was over-intellectualized.[6] When appropriate changes were made at their suggestion, the intelligence test which had contributed heavily to the prediction of the earlier grades dropped to approximately zero predictive efficiency. At other times, administrative changes in the curriculum have been enough to upset the matrix of predictors. Flight training during the war proved to be particularly fluid, involving changes both in the items taught and in the method of appraisal. Such changes always held the threat that we might be predicting what the sponsor desired yesterday rather than what he desired today. Finally, as the nature of the war changed from defensive, through defensive-holding, to

[6] This development took place at a number of centers and under the influence of a number of psychologists. It is to be hoped that an account of their activities will receive publication as a genuine contribution to our knowledge of the psychology of training.

all-out offensive, there was real reason to believe that the qualities desirable in combat aviators were themselves subject to change. Adequately meeting the desires of the sponsor would imply that prediction was adjusted to meet the ultimate changes in the field-situation. Some evidence, presented before the war, suggests that the same may hold true in industry as a given concern shifts over from a period of dynamic expansion to a period of conservative and steady operation.

The materials presented above, even in necessarily abbreviated form, may be enough to indicate increased sophistication in regard to that which psychologists are trying to predict. They may be enough to suggest that the psychologist who readily accepts any convenient measure is guilty of inadequate devotion to his technical responsibility. They should certainly be enough to demonstrate that validation is not a simple technical problem and that the criterion itself may provide the psychologist with as great a technical challenge as the procedures incident to the assembly of good predictors. There is much more to be written on the subject. If this brief paper serves as a stimulus adequate to bring out pertinent discussion from those who have encountered the basic problem in the Armed Services, in industry, and in education, it will have served its purpose.

# ADAPTATION OF THE MINNESOTA MULTIPHASIC PERSONALITY INVENTORY FOR GROUP ADMINISTRATION AND RAPID SCORING[1]

BY ARTHUR BURTON[2] AND CHARLES J. BRIGHT
CALIFORNIA STATE PERSONNEL BOARD AND
INTERNATIONAL BUSINESS MACHINES CORPORATION

THE MINNESOTA Multiphasic Personality Inventory consists of 550 items each on an individual card and covers a range of subject matter from physical condition to social attitudes and occupational choice [1, 2]. In the use of the Inventory as a personnel tool, it was soon recognized that the laborious and time-consuming procedure would be a bar to the wider use of the test as a standard element in a selection program, and generally limit its function as a group test. The experience of clinicians[3] confirmed the findings with respect to personnel work, and the difficulty of obtaining clerical help, coupled with a desire for greater accuracy in scoring, led the writers to seek a more rapid, inexpensive, and accurate means of scoring. In this manner it was believed that the usefulness of the test could be extended both in industry and clinically. Since reformulation of the basic scoring plan seemed inadvisable, it was decided to adapt the test for punched cards. Each item—for example, "During the past year I have been well most of the time"—is sorted by the subject into one of three categories: "True," "False," or "Cannot Say." The sorted responses are then scored for the following symptom-complex scales: Hypochondriasis (Hs), Depression (D), Hysteria (Hy), Psychopathic deviate (Pd), Interest (Mf—masculinity and femininity, Paranoia (Pa), Psychasthenia (Pt), Schizophrenia (Sc), Hypomania (Ma), and Question Score (?), Lie Score (L), and Validity Score (F), validating scales.[4]

The following directions for scoring the Inventory are given by the test authors:

> The first step in recording is to remove the 'True' pack and separate it into two packets of cards, one with the lower right hand and the other with the lower left hand corners cut off. . . . The above separation will divide the 'True' cards into two packs, one consisting of items frequently classified as true by normal persons and one, usually much smaller, composed of items infrequently classified as true. The criterion for frequency, which determined the cutting of the corners, was a tabulation of responses made by the normal subject groups. Only the packet representing the infrequent answers is used in recording scores. Each item in this group is recorded as a red X in the appropriate cell of the record blank identified by code letter and number. After the infrequent 'True' items have been recorded, the items answered as false should be similarly divided into two packets according to the cor-

[1] Since the preparation of this paper, a group form using an alternative technique has been published by the Psychological Corporation.

[2] Now with the California Youth Authority.

[3] Two psychologists in California state institutions reported that the mean scoring time for a skilled clerical worker ranged from 15 to 30 minutes depending upon the nature of the record and the ability of the scorer.

[4] Revised (1943) edition.

ner cuts. In this case the infrequent answers will be those with the lower left hand corners clipped. These items are next entered in the appropriate cells on the record sheet with exactly the same type of red X for the true items. As a final step, the 'Cannot Say' items are recorded without separation in their appropriate cells. . . . The Question score is simply the total number of question marks ['Cannot Say' items]. The L items are the last 15 items in the Inventory. . . . . For the F score, the transparent scoring template marked F is used. The 64 items are printed on the template in proper position with an X so arranged that when the template is placed over the record sheet and the anchor points aligned, the X's will fall beside the recorded items of the scale. To obtain the raw score, count one point for every item having a red X on the record sheet beside an X on the key. . . . . The raw scores for the remaining scales are obtained from the appropriate keys in the manner just described for the F score, except that some keys contain zeros in addition to X's. When a blank cell on the record sheet corresponds to a zero on the key, score one point also for the agreement [1, pp. 7-8].

The adaptation of the Inventory was done in the following manner. A deck of 550 specially designed I.B.M. punched cards was substituted for the original 550 items. Each item was typed on the left hand side of the punched card while the right side contained the ruled card fields used in scoring. The punched card deck was administered to the subjects in exactly the manner as described above for the original deck. That is, the subjects sorted prepunched cards which were used over and over again; the subject's responses were not punched.

A numeric code was assigned to each item as follows:

| Item Classification | Number Code |
|---|---|
| A1 to A55 | 001 to 055 |
| B1 to B55 and so on until | 101 to 155 |
| J1 to J55 | 901 to 955 |

This code is punched in columns 53-55 of the punched card. The equivalent of the card corner cuts and templates (keys) were punched into the card fields for scoring all scales simultaneously as follows:

| Card Field | Card Columns | Test |
|---|---|---|
| A | 56-57 | F |
| B | 58-59 | Hs |
| C | 60-61 | D |
| D | 62-63 | Hy |
| E | 64-65 | Pd |
| F | 66-67 | Mf |
| **Card Field** | **Card Columns** | **Test** |
| G | 68-69 | Pa |
| H | 70-71 | Pt |
| I | 72-73 | Sc |
| J | 74-75 | Ma |
| K | 76-77 | L |

Reference to the punched cards in tabulating would be somewhat simplified by using the scale nomenclature at the head of the card fields, but it was deemed advisable to use the A through K Code to prevent the possibility of giving any lead to the subject during administration of the Inventory.

The problem of scoring resolves itself into counting X's or dissimilar responses (items most frequently classified as true which the subject classifies in the false deck, and vice versa) and also zeros, or similar responses (items classified as true which the subject classifies in the true deck, and vice versa). This is accomplished by punching the digit one in each column of the two-column fields for the scales as required to record the significant responses.

Generally the first column of the two-column fields represents the true statements, and the second column the false. If an item most frequently classified as true should register as an X response if found in the false deck, a digit one is punched in the second column of a field. If an item most frequently classified as false should register as an X response if found in the true deck, a digit one is punched in the first column of a field.

In the case of zero scores: If an item most frequently classified as true should be found in the true deck, a digit one is punched in the first column of the field —and vice versa for the false items. Since cards were to be given the subject in random order, the items were randomized through the use of Selective Service numbers shown by Peatman to have occurred by chance [3].

The plugboard for the I.B.M. tabulator is wired to list and count the digit ones punched in the various columns and to count all cards. The normal method of placing cards in the feed hopper of the machine is face down and nines position first, and the cards selected by the subject as "True" should be placed in the machine in this manner.

Cards classified in the "False" deck should be placed in the machine face *up* and nines first. Cards in the "Cannot Say" deck are placed in the machine either way since only the card count is used in scoring this scale and a total is obtained regardless of the manner in which the cards are placed in the machine.

After wiring the plugboard and before listing the scale score reports, the wiring of the plugboard should be checked by test cards prepared according to general I.B.M. instructions. All machine supervisors and I.B.M. Service Bureau operators can supply this information.

As the cards are received for machine processing, it is recommended that they first be sorted on Code (Cols. 53-55) to simplify reference to a subject's responses, should irregularities appear, or item analysis be desired.

The "True" deck should be sorted in normal manners; i.e., by placing the cards in the sorter face down and nines first, and sorting consecutively on columns 55, 54, and 53.

The "False" deck should be placed in the sorter face *up*, nines first, and sorted consecutively on columns 26, 27, and 28. These columns correspond to 55-53 when the cards are turned over.

The decks are then listed by the I.B.M. tabulator at a speed of 80 cards per minute on the tabulator score sheet (Figure 1). Totals are printed at the bottom of the list. "True" and "False" decks can be listed separately (in which case separate totals appear and must be added together) or consecutively— in which case one total is given. In either case, care must be exercised that the cards are placed in the feed hopper properly to produce the correct scores. The "Cannot Say" deck is run through for a card total.

After completion of the listing on the tabulator score sheet, the decks are all placed in the sorter and sorted together on the Code R, face down and nines first, consecutively on Cols. 80, 79, and 78. This places them in random sequence for administration to the next subject.

Each sorting of the cards and the run through the tabulator takes approximately 7 minutes, or a total of approximately 21 minutes for the processing as outlined above. If there is no need for item analysis or reference to specific items, the preliminary sorting can be dispensed with (saving 7 minutes), and the cards can be tabulated instead of listed. The I.B.M. tabulator tabulates at a rate of 150 cards per minute, which reduces the machine time from 7 minutes to 4 minutes. If there is no desire to present the items to each subject in the same prearranged random sequence, the final sorting can be eliminated—so that the overall machine processing time is cut from 21 minutes to 4 minutes.

The advantages of this method of scoring can now be summarized: Machine accuracy is substituted for human fallibility. The human element is

FIGURE 1. MULTIPHASIC PERSONALITY INVENTORY SCORE SHEET

| Test | F | Hs | D | Hy | Pd | Mf | Pa | Pt | Sc | Ma | L | Item | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Card Code | A | B | C | D | E | F | G | H | I | J | K | Count | |
| 202 | | | | | | | | | | | | | |
| 203 | | | | | | | | | | | | | |
| 205 | | | | | | | | | | | | | |
| 206 | | | | | | | | | | | | | |
| 208 | | | | | | | | | | | | | |
| 209 | | | | | | | | | | | | | |
| 210 | | | | | | | | | | | | | |
| 211 | | | | | | | | | | | | | |
| 213 | 1 | | | | | | | | 1 | | | | |
| 216 | 1 | | | | | | | | 1 | | | | |
| 219 | | | | | | | | | | | | | |
| 221 | | | | | | | | | | | | | |
| 224 | | | | | | | | | | | | | |
| 225 | | | | 1 | | | | | | | | | |
| 227 | | | | | | | | | | 1 | | | |
| 228 | | | | | | | | | | | | | |
| 231 | | | 1 | | | | | | | | | | |
| 232 | | | | | | | | | | | | | |
| 233 | | | 1 | 1 | | | | | | | | | |
| 234 | | | | | | | | | | | | | |
| 236 | | | | | | | | | | | | | |
| 244 | | | | | | | | | | | | | |
| 246 | | | | | | | | | | | | | |
| 247 | 1 | | | | 1 | | | | 1 | | | | |
| 250 | | | | | | | | | | | | | |
| 251 | | | | | | | | | | | | | |
| 253 | | | | | | | | | | | | | |
| 254 | | | | | | | | | | | | | |
| 255 | | | | | | | | | | | | | |
| 303 | | | | | | | | | | | | | |
| 304 | | | | | | | | | | | | | |
| 305 | | | | | | | | | | | | | |
| 308 | | | | | | | | | | | | | |
| 310 | | | 1 | | | | | | | | | | |
| 311 | | | | | | | | | | | | | |
| 312 | | | | | | | | | | | | | |
| 313 | | | | | | | | | | | | | |
| 316 | | | | | | | | | | | | | |
| 317 | 1 | | | | | | | | | | | | |
| Total | 4 | | 3 | 2 | 1 | | | | 3 | 1 | | 39 | * |

reduced to a minimum since all of the basic scoring information is punched into the deck of cards and new decks can be reproduced mechanically from the original deck. The possibility of error in the original punching is not great. All possibilities of error are offset by safeguards in procedure. Machine wiring is checked in advance by test cards. Overall check of cards is maintained since total card count of all decks must equal 550 and this count is automatically obtained.

Possibility of intermingling true and false decks is reduced to a minimum, since they should be processed sepa-

rately.

The principal possibility of error lies in putting the cards in the tabulator incorrectly. This too can be checked without readministering the test for usually the false deck is larger than the true deck and through observation of the total card count, it can be determined how the cards were placed in the machine. Since the code numbers are listed, it is possible in such case to reconstruct the test and score it properly.

All tests are scored simultaneously and at from 80 to 150 cards per minute. For mass scoring, e.g., in a personnel selection program, scoring time can be reduced to 4 minutes a test, and by elimination of non-diagnostic items, the scoring time can be reduced below 4 minutes. There is no fatigue factor with a machine; it delivers the same speed and accuracy after 7 hours of constant work as after 7 minutes.

An uncontrolled factor in the Inventory is that responses to certain statements might vary depending on the sequence of presentation. The prearranged random code insures that all subjects start with decks of cards in identical order. This is not provided for in the original Inventory.

Personnel agencies and state, county, or other institutions may have I.B.M. machines of their own and in these cases scoring of the inventory represents no additional financial outlay for machines and probably little for operator hire.

For those without I.B.M. machines, the I.B.M. Service Bureaus are available in all principal cities of the United States, Canada, South America, and most United Nations countries. In these I.B.M. offices, I.B.M. machines and experienced personnel are available on an hourly basis for work of just this sort. The cost of this service is moderate and less than the cost of manual scoring where a considerable number of tests are to be scored.

The increased rapidity of scoring and the possibility of reproducing the decks mechanically by I.B.M. equipment—where card forms have been printed—makes it possible to administer this Inventory as a group test.[5] It also makes validation of existing scales and the formulation of new scales easier.

The procedure outlined here would seem to make the Minnesota Multiphasic Personality Inventory adaptable for wider usage by school systems, personnel agencies, correctional schools, mental hospitals, and others.

[5] The purpose of this artice is essentially heuristic. The Inventory, of course, cannot be reproduced without permission of the copyright owners. It is hoped that the publishers will issue a punched card edition of the Inventory as illustrated here and permit duplication of the decks on a fee basis, or issue prepunched card decks at a nominal charge.

## References

1. Hathaway, S. R. and McKinley, J. C. *The Minnesota Multiphasic Personality Inventory.* Minneapolis: University of Minnesota Press, 1943. Pp. 16.
2. ———. "A Multiphasic Personality Schedule (Minnesota): I, Construction of the Schedule," *Journal of Psychology,* 1940, 10: 249-54.
3. Peatman, J. G. and Schafer, R. "A Table of Random Numbers from Selective Service Numbers," *Journal of Psychology,* 1942, 10: 295-305.

# PROFESSIONAL STATUS AND TRAINING OF PSYCHOLOGISTS

## REPORT ON ANNUAL MEETING OF NYSAAP

BY LOUIS LONG

STUDENT PERSONNEL BUREAU, COLLEGE OF THE CITY OF NEW YORK

THE NINTH Annual Meeting of the New York State Association for Applied Psychology, held at Albany, New York on January 26, 1946 was devoted largely to the consideration of professional problems, such as certification of psychologists, internships for students, agencies utilizing psychologists and misrepresentation in the field of psychological guidance.

### A. PROPOSED CERTIFICATION PROGRAM

The members of the NYSAAP unanimously voted to instruct the Committee on Certification[1] to submit its recommendations to the membership for action by mail ballot. The program of certification will be adopted if approved by two-thirds of the votes returned. The Committee has recommended that the Association set up its own machinery for certification since it was felt that it would not be feasible at this time to attempt legal certification, although this is the ultimate goal of the certification program. More experience with the problems of certification is necessary before any type of legislation can be sought.

The plan recommended calls for the certification of psychologists doing work in the clinical field, which for purposes of certification the Committee has defined to include personnel work, counseling, the work of the school psychologist, as well as other phases of clinical work. The Committee's plan is presented below.

1. Board of Certification

There shall be established a Board of Certification. This Board shall consist of five members nominated by the Executive Committee of the NYSAAP and elected by the membership for a period not to exceed three years. There shall be two names submitted for each vacancy. The membership shall also have the right to nominate by petition. The term of office of the members of the Board shall be staggered so as to insure continuity of office. The Secretary of the Association shall serve as an ex officio member of the Board, but without a vote on this Board.

The Board of Certification shall report to the Executive Committee not less than twice a year and to the membership at large at the annual meeting of the Association.

The Board of Certification shall be empowered to grant a Certificate in Clinical Psychology to all persons who meet the requirements set forth below, upon application and the payment of a fee of five dollars.

2. Professional Requirements for Certification

The professional requirements for a Certificate in Clinical Psychology are:

*a*) A Ph.D. degree in psychology or

[1] Members of the Committee are: Mr. Simon H. Tulchin, *Chairman*, Dr. Morris Krugman, Dr. Louis Long and Dr. J. Wayne Wrightstone.

applied psychology, or the equivalent; one year of experience in the application of clinical psychology under the direction of a psychologist with the qualifications of a Fellow in the American Psychological Association; and, in addition, one year of experience in the clinical field; or

*b*) An M.A. degree in psychology, or the equivalent; and four years of experience in the application of psychology, two of which shall have been in the application of clinical psychology, and one of these two years shall have been under the direction of a psychologist with the qualifications of a Fellow in the American Psychological Association; or

*c*) Training and experience considered by two thirds of the members of the Executive Committee to be equivalent to the requirements set forth in (*a*) or (*b*).

3. Other Requirements for Certification

Applicants for the Certificate in Clinical Psychology are required to reside or be employed in New York State. They must meet ethical and professional standards set up by the NYSAAP and must be sponsored by two persons who are members of the NYSAAP or Fellows of the APA. These sponsors should be familiar with the applicant's ethical and professional work standards. Applicants are *not* required to be members of the NYSAAP.

4. Revoking a Certificate

A certificate may be suspended or revoked for a definite period of time for good and sufficient reason, after a due hearing and upon approval of the Executive Committee.

5. Review of Program of Certification

An evaluation of the functions and the fees of the Board of Certification shall be made by the Executive Committee from time to time and the Executive Committee shall have the power to modify any of the provisions presented above, subject to the approval of the membership at the Annual Business Meeting of the Association.

It should be noted that the requirements for certification differ only slightly from the requirements for membership in the NYSAAP. For membership, persons with a Ph.D. need one year of "supervised" experience in the application of psychology; whereas, for certification, they need one year of "supervised" experience in the application of *clinical* psychology, plus one year of additional experience in the clinical field. For membership, persons with an M.A. need four years of experience, one of which shall have been "supervised"; whereas, for certification, they need the same number of years of experience, but during this period they must have had two years of experience in the *clinical* field, one of which shall have been "supervised."

The Executive Committee would be given the power to approve the certifying of well-qualified individuals who do not, for one reason or another, meet the requirements listed in (*a*) or (*b*) of Section 2. Similar action can now be taken by the Executive Committee in granting membership. Some such provision is necessary if injustices are to be avoided and this is especially true when a new program is being established.

The members of the Committee on Certification emphasized the fact that although the process of certification will in itself further the standing of the psychologist as a professional person, an active attempt must be made to educate the public and the employing agencies so that both will more readily accept psychology as a profession. As the professional status becomes better estab-

lished it will be easier to bring about improvements in training facilities and eventually this should lead to the development of professional schools for psychologists.

### B. INTERN TRAINING

The need for the improvement of training facilities for students of psychology was also stressed by the Committee on Psychological Internships.[2] With the aim of encouraging educational and training institutions to set up additional training programs in clinical psychology this Committee formulated a definition of psychological internship which considers such questions as preinternship training, duration of internship, and adequacy of supervision. The definition under consideration by the NYSAAP is as follows:

> A psychological internship shall be considered a training position for persons with a Master's degree in psychology, or the equivalent, requiring full-time work for a period of nine months or longer (but not over two years) under adequate supervision. Adequate supervision shall be construed to mean supervision by a psychologist of recognized professional standing, whose training and experience meet any one of the following, or their equivalent: (1) the requirements of the New York State Department of Civil Service for the Psychologist or Senior Psychologist grade, (2) the requirements for fellowship in the American Psychological Association, or (3) requirements as formulated and approved by the New York State Association for Applied Psychology; and in addition whose program of work allows sufficient time for supervision of interns.

As a preliminary step in the investigation of opportunities for internship, letters were sent to 17 members of the NYSAAP who were associated with organizations doing clinical work or who were members of the Psychology Department of various colleges or universities. These persons were selected because it was believed that their institutions might either have an intern training program already or be in a position to establish one. Replies were received from 15 of the 17 members.

Six replies discussed internships for school psychologists but among these only the Bureau of Child Guidance of New York City actually had such positions at the present time. Members of the New York State Department of Education expressed active interest in the development of an internship program for school psychologists. One school representative said that internships had been offered in earlier years and probably would be offered in the future. Another respondent stated that such a training program could not be considered because of the heavy schedule of work already carried by a limited staff.

Replies were received from two University Clinics. One of the two has an internship program, whereas the other utilizes graduate assistants and has facilities for supervising training, but has no budget provision for an intern.

Replies were received from the psychology departments of two hospitals. One has internships and the other hopes to establish such positions in the near future.

Of seven agencies and clinics that were represented, only two have established training positions. Of the remaining five, one used to have an internship program and is hopeful about its re-establishment. Two other agencies reported that they will ultimately have training positions. The reports from the remaining two indicate that present conditions are such that internship training could not be recommended.

It seems evident from this preliminary survey that the greatest present need is for the development of intern-

[2] Members of the Committee are: Dr. Elaine F. Kinder, *Chairman*, Dr. Paul S. Achilles, Dr. Henry E. Garrett, Dr. Marion P. Jenkins, Dr. Louis Long, and Miss Gladys Tallman.

ships for school psychologists. It is appropriate at this point to anticipate the discussion of the next section to the extent of saying that the data collected during a more extensive survey by the Association's Committee on Clinics support this conclusion.

### C. AGENCIES UTILIZING PSYCHOLOGISTS

The Committee on Clinics[3] made an exploratory survey of psychological resources in New York State by sending questionnaires to 21 school systems, 26 institutions for mental defectives, 20 state and federal hospitals for mental disease, 26 licensed institutions for mental disease, 14 general hospitals with mental hygiene clinics, 14 community welfare agencies, 10 vocational guidance agencies, and 4 university guidance centers; a total of 135 agencies or institutions. The 58 agencies which responded reported a total of 141 psychologists on their staffs. This does not include unpaid graduate students engaged in psychological work done in connection with course requirements. Of the 141 psychologists, 108 were full-time, paid staff members; 24 part-time, paid staff members; and 9 unpaid staff volunteers. Only 17 of the entire number occupied internship positions; however, several agencies that have internship programs did not report until after the tabulations were made. Consequently the number of such positions is somewhat greater than 17.

1. Details Concerning internships

Nine of the 10 school systems that responded have a psychological staff, but only 2 provide for training positions (White Plains which has a full-time, paid assistant to the school psychologist and New York City which has facilities for one student psychologist). Of the 11 institutions for mental defectives that reported, 9 have a psychological staff, but only one (Letchworth Village) provides for internships. Of the 16 hospitals for mental disease that responded, 5 have a psychological staff, but only 1 provides for internships (two graduate students employed as psychological interns with practically no supervision). Of the 6 hospitals with mental hygiene clinics that reported, each has a psychological staff and three provide for internships (Presbyterian, Bellevue, and Jewish Hospitals). Of the 7 welfare agencies that reported, 5 have a psychological staff, but only one (Brooklyn Guidance Center) has an internship program. Of the 7 vocational guidance agencies that responded, 4 have a psychological staff, but only one (YMCA Vocational Service Center) has an internship program. Of the 3 university guidance centers that reported, all provide for graduate students in training and one provides for a paid training fellowship of the internship type.

2. Distribution of Psychologists

Some idea of the distribution of the 141 psychologists employed by the 58 agencies that responded can be obtained from Table 1. It should be pointed out,

TABLE 1.

DISTRIBUTION OF PSYCHOLOGISTS

| Agency | No. Institutions Replying | No. Psychologists Employed |
|---|---|---|
| School systems | 10 | 51 |
| Institutions for mentally defective | 11 | 12 |
| State and federal hospitals | 16 | 8 |
| General hospitals with mental hygiene clinics | 6 | 23 |
| Welfare agencies | 7 | 10 |
| Vocational guidance agencies | 5 | 15 |
| University guidance centers | 3 | 22 |

[3] Members of the Committee are: Dr. Harold Abelson, *Chairman*, Dr. Ethel Cornell, and Dr. David Wechsler.

however, that psychologists are not evenly distributed among the various agencies of any one type. For example, of the 51 psychologists employed in school systems, 38 are in one system, 3 in another, 2 in each of three other systems, 1 in each of four others; one school system reporting had no psychologist. It was noted that the majority of the agencies employing psychologists have only one such person; consequently, it is unlikely that the interns will have an opportunity to work with more than one psychologist.

### D. MISREPRESENTATION IN THE FIELD OF PSYCHOLOGICAL GUIDANCE

The importance of certification was also emphasized by the members of the Committee on Guidance Practices.[4] This committee feels keenly the need for a list of approved agencies and individual psychologists which could be presented to the public. It is believed that a helpful step in this direction would be the establishment of the plan for certification.

During the last year this Committee prepared a statement on the principles concerning the advertising of aptitude tests and psychological guidance and copies were mailed to all of the New York City newspapers, the telephone company, certain magazines, and to the Committee on Ethics of the APA (Offprints of this statement may be obtained from Dr. Albert J. Harris of the College of the City of New York.)

[4] Members of the Committee are: Dr. Albert J. Harris, *Chairman*, Dr. Warren G. Findley, Dr. Joseph V. Hanna, Dr. Nina Ridenour, and Dr. Harold C. Seymour.

The Committee has received assurances from the *New York Herald-Tribune*, the *World-Telegram*, and the *Sun* of their desire to comply with the principles approved by the Association. Certain advertisements appearing in the *New York Times* and the *Post*, to which objections were raised, have not reappeared. The members of the Committee have asked the telephone company to delete some listings and to modify others. In the last analysis, however, it is apparent that it will be impossible to prevent poorly qualified people from advertising themselves as psychologists until a system of legal certification of psychologists is established.

A study has been made by the Committee to determine what proportion of the people who present themselves to the public as psychologists or vocational advisers have any professional affiliations through which pressure in the direction of ethical practices could be brought to bear, without legal sanctions. The names of 42 individuals were found under the heading of "Psychologists" or "Vocational Guidance" in the *Classified Telephone Directories* of New York State. These names were checked against the membership lists of the APA, the AAAP, the NYSAAP, and the list of Qualified Psychologists in New York State. Only eight of the 42 names appear on any of these lists.

Plans are being made by the Committee to investigate the 39 agencies or organizations that advertise psychological or vocational services. It is hoped that other psychologists will volunteer to help in this important work.

# ON CERTAIN MISUSES OF THE WECHSLER-BELLEVUE SCALES

By A. S. LUCHINS
YESHIVA UNIVERSITY

IN THE PAST three years the writer has read over 2,000 reports of Wechsler-Bellevue examinations written by some 50 different clinical psychologists. Recurrent through them were certain misuses and misinterpretations. In view of the number of psychologists and psychiatrists involved it is important to consider their wrong practices and to discuss means of correcting them.

In most reports the Wechsler IQ's were interpreted on the basis of Terman's or some other "official" classification. Such a procedure is specifically prohibited by Wechsler: "The examiner . . . may not use another scheme of interpreting these same IQ's, for example, that of Terman or Kuhlman . . . he is not privileged to disregard the IQ interpretation as given" [2, p. 191]. Wechsler's insistence on the use of his scheme of IQ interpretation is based on theoretical, practical and statistical considerations clearly presented in the fourth chapter of his manual. These reasons are understandable to those who have knowledge of elementary principles of mental measurement.

In defending this practice many of the psychologists and psychiatrists claimed that they were accustomed to a certain scheme and wanted the test interpreted in accordance with it. What was forgotten was that the scheme they selected might have been proper for IQ's derived from one test but improper for those derived from another test. Each test should be interpreted in accordance with its standards in spite of tradition, custom, or habit of thinking. Neither personal likes nor habits can be substituted for prescribed standards.

Another prevalent practice was the categorizing of an individual's intelligence on the basis of the MA obtained in the Wechsler manual; but the scheme of MA classification was some "official" arrangement of MA's. For example, all MA's below 10 were designated as belonging to the Mental Defective Category of Adult Intelligence. Accordingly, a weighted score of 43 which yields an MA of 9 years was interpreted as placing the adult in the Mental Defective Category. Yet this same score yields for a 19 year old adult an IQ of 58, but for a 36-year-old adult an IQ of 71. The former IQ when properly interpreted places the 19-year-old adult in the Mental Defective Category but the 36-year-old adult in the Borderline Category. Similarly, a score which yields an MA of 10 years might give an IQ which may place one in the Borderline or the Dull Normal Category of Adult Intelligence depending upon his age. It is fairly obvious that an MA means a different thing if it is obtained by a 6-, 10-, 20-, 30-, or 60-year-old individual. An MA does not represent an absolute power of functioning or ability and cannot be used by itself to categorize an individual's intelligence. To ignore the individual's age and not to

avail oneself of Wechsler's IQ tables is to ignore one of the most valuable features of the test. The interpretation of the weighted scores by Wechsler purposely takes into account the rise and decline of intellectual growth with age.

This use of the MA was often justified on the same grounds given for the misinterpretation of the IQ's or by the argument that "after all, the patient does have the level of intelligence represented by the MA," neither of them valid reasons.

"This IQ is a true reflection of the patient's intelligence," or "This IQ is valid," was found in the conclusion of many reports even when the patient was an illiterate, or came from such an educational and cultural background that he was not acquainted with the tasks of the test or had had no opportunity to learn the knowledge and skills needed in their performance. When asked the meaning of these phrases, some psychologists stated that since the test was correctly administered, since there was an absence of symptoms indicating a psychosis or psychoneurosis, and no physical defects which hindered his functioning, etc., the obtained score gave a true index of intelligence. This indicates a confusion in thinking about validity and reliability, concepts which are usually taught in elementary courses in testing.

Some psychologists and psychiatrists defended such conclusions by claiming that they needed, and the test score gave, an index of what the person could do. If the test showed that he was functioning on the mental defective level, then for all practical purposes he was a mental defective. A few elaborated on this argument by stating that as far as the institution which needed his service was concerned, the individual was a mental defective. What is ignored by this point of view is that a person who had no chance to learn and therefore cannot answer the test is not equivalent to a person who had the opportunity and did not learn. It is possible for an institution to set up standards for recruiting of personnel and to exclude all who do not meet them. But it is not proper for that institution's psychologists and psychiatrists to say that regardless of the reason why one cannot meet the standards one is a mental defective. One may be deficient in the trait needed by the institution but not be mentally defective.

Not infrequent were statements similar to these: "This IQ is in keeping with his educational achievement," or "His occupational history reveals that his IQ is valid." Although it is possible that one is a laborer or never went beyond the second grade in school because of his ability, it is also possible that because of his educational and vocational experience he scored low on the test. This type of interpretation has little or no basis in the current literature on mental measurement.

That most reports did not merely give an IQ but contained attempts to analyze the nature of the scatter of the subtest scores showed that the psychologists were up to date on the latest practices, but in many instances the interpretations were forced. Diagnoses were made on the basis of the scatter in spite of the fact that only three verbal and three performance tests were used with some of the important subtests omitted. For instance, Object Assembly and Similarities were missing from a battery of tests given to determine organic involvement. In some instances the reports stated that the scatter was, for such a per cent, equivalent to this or that scatter found in Wechsler's or Rapaport's manual [1], e.g., 30 per cent psychotic and 70 per cent psychoneurotic. Such practices are not warranted

by the present knowledge of scatter analysis.

An IQ can be approximated by the use of less than the entire battery, but not less than eight tests should be employed [2, p. 171]. Not infrequently only five or six tests were used to obtain an IQ. Moreover, these were often selected without reference to the patient's educational and cultural background, e.g., a rural southern colored man who had no schooling was given a battery of six tests, four of which involved school-learned matter: Vocabulary, Similarities, Arithmetic and Digit Symbols.

When these findings were discussed with some of the clinical psychologists, a few were genuinely surprised to learn that they had incorrectly interpreted the IQ's or MA's. They had actually not read the related discussions in Wechsler's manual. Some claimed that the psychiatrists insisted on such interpretations; still others stated that psychologists with graduate degrees, under whom they worked, taught them to do what they did. One may dismiss the above practices as not being the work of "real" clinical psychologists but incompetents. In many cases, however, charges of incompetency are definitely out of place. Something besides name calling is needed. The profession owes it to the public and its own development as a clinical art to consider methods of preventing such misuse of tests.

As a beginning in remedying this state of affairs, the writer proposes the following:

(1) A committee be created to set up standard operating procedures for interpretation of various clinical instruments with special emphasis on pitfalls to be avoided.

(2) Journals devoted to practical, applied, consulting and clinical psychology should run frequent reviews of basic principles, as well as new research in the field of testing. Also, they should run a "boner" department.

(3) Liaison should be established with the American Psychiatric Association so that its members will be made cognizant of acceptable practices in clinical psychology. In some instances the writer has found psychiatrists to be ignorant of the basic principles of mental measurement or of recent developments in the field. Yet some of these men set up standards and insisted upon certain tests being used for functions for which they were not constructed. Clinical psychologists cannot function in an atmosphere of ignorance.

REFERENCES

1. Rapaport, D., with collaboration of Schafer, R., and Gill, M. Manual of Diagnostic Psychological Testing. 1. Diagnostic Testing of Intelligence and Concept Formation. *Josiah Macy, Jr. Foundation Review Series*, 1944, 2, No. 2.
2. Rapaport, D. *Diagnostic Psychological Testing.* Chicago: The Year Book Publishers Inc., 1945. Vol. 1.
3. Wechsler. D. *The Measurement of Adult Intelligence.* (3rd Ed.) Baltimore: Williams & Wilkins, 1944.

# BOOK REVIEWS

**HUMAN RELATIONS IN INDUSTRY.** By Burleigh B. Gardner. Chicago: Richard D. Irwin, Inc., 1945. Pp. xi + 307. $3.75.

This graphic and realistic picture of human relations in industry is not a prosaic description but an effective presentation that holds interest to the end. The reader is given insight into the social structure and psychological interaction of real people in industry as the basis for understanding many of the problems and stresses in employee-management relations. The human element is so real that any reader who has worked in an industrial shop or office will recall with a smile or chuckle experiences that fit into the author's picture.

The first three chapters describe the factory as a social system, the lines of authority and communication, and the functions and problems at each level. Here we see the great importance of status to the members of this social structure. Status is derived from position in the supervisory hierarchy, from shop-office distinctions, and from other symbols of status. Added to the formal organization are many informal but effective relationships which "place" people and provide grapevine lines of communication.

Chapter IV on segmentation of the structure describes the cliques which arise, the horizontal cleavage and the relationships of line and staff. Chapter V on the union and its functions and place in the structure indicates a needed relationship and line of communication provided by the union that is seldom adequate in the formal organization. The next two chapters discuss wage systems and their inadequacy as incentives.

Chapter VIII, the individual in the structure, Chapter IX, the industrial relations organization, and Chapter X, Personnel Counseling, deal with the problems of adjustment of the individual in the complex of human relations. Included are the non-directive techniques of counseling developed in the Western Electric Company and also independently by Carl Rogers.

Chapter XI, on minority groups in industry, describes the caste system and the social difficulties of bringing into an industry any number of foreigners, negroes, or women. The next chapter on problems of co-operation, and the final chapter on understanding social structure, indicate the weak points in co-operation in most factories and present the understandings of social structure as tools for remedying these deficiencies. Much of these chapters, in fact of the entire book, seems elementary or common-sense, but with experience in industry, one learns there are many executives who consider this human element interesting but unimportant, or are very naïve about it all.

Although there is no complete substitute for work experience for guidance, personnel or social workers, this book will go far in helping to give an appreciation of human relations as an important factor in vocational adjustment.

Certainly this adjustment depends on more than placing square pegs in square holes. More than techniques of testing, or a knowledge of occupations gained from reading job descriptions is needed for adequate insight into personnel or counseling work. In fact, all groups in industry, even the experienced, have frequently lacked the objective insight to analyze and cope with the social structure and psychological forces in which they worked. This very readable book can well be used in training and discussion groups of foremen and others higher in the hierarchy of the vocational world.

PROFESSOR BRUCE V. MOORE
THE PENNSYLVANIA
STATE COLLEGE

**WHAT PEOPLE ARE.** By Clark W. Heath et alli. Cambridge, Mass.: Harvard University Press, 1945. $2.00.

**YOUNG MAN, YOU ARE NORMAL.** By Earnest A. Hooton. New York: G. P. Putnam's Sons, 1945. $2.50.

These volumes, the first general publications on the Grant Study, report a research undertaking developed as a result of a hygiene professor's conviction that we do not know enough about human nature to guide young men wisely

toward full utilization of their potentialities. Since 1938, Dr. Arlie Bock has implemented his idea through the study of 268 "normal" undergraduates by a staff composed of two psychiatrists, a physician, a physiologist, a physical anthropologist, a psychometrist, and a "personnel worker," the latter being responsible for material on socio-economic status and family background. The immediate goals have been to describe the group in terms of the separate approaches represented on the staff and also to determine the interrelationships among variables measured by the different techniques. To date, they have compiled many population statistics and recorded the associations among characteristics by the degree of overlap. While it is impossible to summarize the undigested mass of specific findings, two examples may be mentioned. They report what traits are associated with "soundness" of personality and with various types of personality integration. They have found that the masculine component in physique is more closely related to psychological traits and other characteristics than are somatotype and bodily disproportions.

The Grant Study has the merit of bringing together representatives of different disciplines for a common goal, the study of human behavior. Moreover, this "institution" plans to follow its subjects over a period of years. A further distinctive characteristic is the absence of preconceived theoretical bias. On the other hand, the potential value of the Study is unfortunately offset by such serious liabilities as inadequate quantification of data and unwillingness to draw upon the experience of others. The psychiatrists preferred to approach the problem of describing personality *de novo*, instead of consulting the vast amount of relevant psychological literature. There is little evidence of any familiarity with similar types of projects, including the one conducted in the geographical vicinity of the Grant Study, reported in "Physique, Personality and Scholarship."

Probably because of their clinical approach, these investigators began by grouping individuals who showed a marked degree of a given trait. But the dividing line is not sharply defined and at best they are left with a crude two-step scale. The problem of correlation is handled by descriptive percentages with no explicit statistical analyses. One might infer that their desire to be free from preconceived frameworks has prevented them from collecting data in a form suitable for accurately determining degrees of relationships.

Heath's volume outlines the aims and methods of the Grant Study. Writing for the trained worker, he gives much of the technical description needed for an evaluation of the findings but he includes only the obtained distributions of selected characteristics and a few interrelationships. The reader gains a clear conception of the staff's philosophy and techniques but is left with an unsatisfactory answer to his natural inquiry into the Study's contributions to the understanding of human behavior. The meager quantitative data presented by Heath are in terms of percentage distributions for pairs of variables. No statistical measures of degree or significance of relationship are used, except for a ratio between percentages of extreme groups possessing a given trait—a questionable statistic.

To present their findings to the lay reader, the staff obtained the services of Hooton and were rewarded by a thoroughly readable book written in his usual vigorous, provocative style. The material covered by Hooton includes not only a summary of Heath's content but also a detailed account of obtained associations presented without tables. By reconstructing scatter-diagrams and setting up distributions, the painstaking reader can assay the findings for himself, but the lay reader will probably accept the stated conclusions without concerning himself with Hooton's generalized statements on the reliability of the data. The reviewer's curiosity prompted him to evaluate the statistical significance of certain differences in proportions; these were found to be unreliable in spite of Hooton's opinion that the "correlation . . . . seems to be significant." While many suggestive hypotheses can be gleaned from this report, one cannot help wondering why Hooton was allowed to interpret for the layman a mass of data on interrelationships which Heath excludes from his careful technical book. In fact, both authors suggest that it is too early yet for the Study to present results.

Since they have made their reports, it is not too early for a critical examination of the entire effort, especially since it is to be continued. In spite of the emphasis upon a clinical approach and the absence of group testing, there is little attention paid to individuals in either of the published volumes, the exception being the case summaries in the section on mental measurement. Yet this very emphasis made the staff members unprepared for the essential problem of studying interrelationships. While they say that their inhibited subjects are flat-chested (because 33 per cent of

those labelled "inhibited" meet a certain criterion of body proportions), they are unable to tell us from their percentages whether this "inhibited" characteristic is associated more with flat-chestedness or with soundness of personality, and they cannot state whether it is significantly related to either.

Perhaps in its own independent way the Grant Study will eventually reach its objective by gaining an understanding of human behavior which may lead to great advances in vocational guidance. Little in either book suggests that it has made much progress to date.

DONALD W. FISKE

NAVY DEPARTMENT
WASHINGTON, D.C.

PERMANENT LEARNING. By W. H. Lancelot. New York: John Wiley and Sons, 1944. Pp. X + 221.

*Permanent Learning* grew out of another volume by the same author, *Handbook of Teaching Skills*, and partakes of the nature of its predecessor. It is essentially a handbook bringing together as it does guiding principles of learning and presenting them in a compact, serial form. The learning to be made permanent is made up of truths which are essential to life and which contribute to the realization of the seven cardinal aims of Education. For these truths and principles to become permanent pupils must be interested in them, must understand them, and must use them frequently.

In order that interest be aroused twelve procedures are set down, explained and then summarized in a compact, serial form of twelve rules to be followed. Three of the twelve rules are (1) appeal to natural impulses (2) select knowledge that affects the pupils themselves, those about them, or humanity at large and (3) inform them of their progress. In like manner lists of principles are formulated to lead pupils to adopt certain ideals, to secure clear understanding, and for acquiring needed abilities. The needed abilities consist of ability to think well, to reason exactly, to develop judgment, to cultivate one's creative abilities, and to acquire facility in one's manipulative abilities. The outstanding emphasis in every one of these procedures is placed upon interest and understanding, and upon thinking out problems for oneself. The author points out that advance in understanding is not always secured by a too great reliance upon individual problems unless they contribute to the larger understanding and insists that the good teacher must be an arouser and stimulator of those interests and understandings which are essential. The text offers suggestions for selecting good problems and for preparing them for class discussion. In the final chapter all procedures are summarized in a form suitable for quick reference.

There is certainly no quarrel with the selection of principles applicable to the processes of making learning permanent nor with the attractive printing and careful editing. The reviewer believes that such handbooks are wrong in principle. The principle of construction is wrong because the serializing of a large number of rules or principles encourages memorizing by the teacher without complete understanding. A few broad principles thoroughly understood because their meaning has been made clear through experimentation, logic, and numerous illustrations are more apt to find their way into practice than a larger number of rules neatly set forth in serial order. To be concrete, the understanding of interest in general and of children's interests in particular more nearly comes in the examination and study of lists of books and magazines, movies, and radio programs than in the listing of twelve rules to be followed in its attainment.

A. M. JORDAN

UNIVERSITY OF
NORTH CAROLINA